

PROTEOPOLYMERSOME: A VERSATILE TOOL TO STUDY MICROSOMAL MONOOXYGENASES AND FOR DRUG SCREENING

**By:**

Hossam Omar Ali

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Faculty of Engineering

Department of Chemical and Biological Engineering

August 2017

**Acknowledgements**

I would like to thank The University of Sheffield, and the Institute of Materials Research and Engineering at A\*STAR for offering me this valuable experience and for their financial support.

I would like to thank my supervisors Professor Nallani and Dr Wong for their continuous guidance and support, their mentoring was invaluable in getting me this far. I have been very fortunate to have two great supervisors who were always supportive whenever I needed a chat. I have learnt a lot from you both, inside the lab and outside of it, all of it I’m sure will be very valuable in my future. I would also like to thank Dr O’Shea for accepting me as a student at IMRE. I would like to show my appreciation to all the staff, both academic and administrative, at the University of Sheffield, at CBSS in Nanyang Technological University and at IMRE, you have made this Ph.D. a wonderful learning experience.

Everyone with Tuck’s group; Pawel, Zaki, Inas, Abdul, Miriam and Yomi, and everyone at CBSS and ACM; Tom, Fabian, James, Zhaolong, Kelvin and William I would like to thank you for all your amazing help and support, I have been very fortunate to share lab space with you all. I would like to express special appreciation and thanks to Abbas, you have helped me settle in Singapore and becoming as great a friend as I could wish for, again thank you for helping me with my research work, you were always willing to lend a helping hand and make the polymersomes I needed.

I would like to thank my brothers and all my friends for their support, patience and friendship throughout my Ph.D. Finally, I would like to offer a special acknowledgement to my Mum and Dad, for all the love, support and guidance over the years, I can never thank you enough.

Table of Contents

[Abstract - 1 -](#_Toc504580815)

[1. Literature Review - 2 -](#_Toc504580816)

[1.1. Summary - 4 -](#_Toc504580817)

[1.2. Microsomal monooxygenases - 7 -](#_Toc504580818)

[1.2.1. Introduction - 7 -](#_Toc504580819)

[1.2.2. Naming and classification - 8 -](#_Toc504580820)

[1.2.3. Microsomal monooxygenases - 10 -](#_Toc504580821)

[1.2.4. Cytochrome P450 3A4 - 11 -](#_Toc504580822)

[1.2.5. Cytochrome P450 reductase - 15 -](#_Toc504580823)

[1.2.6. Cytochrome *b5* - 16 -](#_Toc504580824)

[1.2.7. Enzyme engineering - 19 -](#_Toc504580825)

[1.2.8. Future directions and research - 25 -](#_Toc504580826)

[1.3. Cell-free protein expression - 26 -](#_Toc504580827)

[1.3.1. Introduction - 26 -](#_Toc504580828)

[1.3.2. Cell-free expression - 27 -](#_Toc504580829)

[1.3.3. Extracts and lysates - 29 -](#_Toc504580830)

[1.3.4. Vectors - 37 -](#_Toc504580831)

[1.3.5. Energy sources - 41 -](#_Toc504580832)

[1.3.6. Post-translational modifications - 43 -](#_Toc504580833)

[1.3.7. Commercial CF systems - 47 -](#_Toc504580834)

[1.3.8. Applications - 49 -](#_Toc504580835)

[1.3.9. Future outlooks - 58 -](#_Toc504580836)

[1.4. Biomimetic cell membranes - 60 -](#_Toc504580837)

[1.4.1. Introduction - 60 -](#_Toc504580838)

[1.4.2. Biomimetic membranes - 61 -](#_Toc504580839)

[1.4.3. Polymersomes - 63 -](#_Toc504580840)

[1.4.4. Proteopolymersomes - 66 -](#_Toc504580841)

[1.4.5. Targeted polymersome insertion - 71 -](#_Toc504580842)

[1.4.6. Future perspectives - 72 -](#_Toc504580843)

[1.5. Aims - 74 -](#_Toc504580844)

[2. Methodology - 75 -](#_Toc504580845)

[2.1. Materials - 77 -](#_Toc504580846)

[2.2. Molecular cloning - 78 -](#_Toc504580847)

[2.2.1. CYP3A4 - 78 -](#_Toc504580848)

[2.2.2. CPR - 81 -](#_Toc504580849)

[2.2.3. C*b5* - 82 -](#_Toc504580850)

[2.2.4. Membrane peptides - 82 -](#_Toc504580851)

[2.3. Vesicle preparation - 85 -](#_Toc504580852)

[2.3.1. Polymersomes - 85 -](#_Toc504580853)

[2.3.2. Liposomes - 87 -](#_Toc504580854)

[2.4. Protein expression - 89 -](#_Toc504580855)

[*2.4.1.* *E. coli* - 89 -](#_Toc504580856)

[2.4.2. Insect CF - 89 -](#_Toc504580857)

[2.4.3. HeLa CF - 90 -](#_Toc504580858)

[2.4.4. Wheat germ CF - 90 -](#_Toc504580859)

[2.5. Analysis of protein expression - 91 -](#_Toc504580860)

[2.5.1. SDS-PAGE - 91 -](#_Toc504580861)

[2.5.2. Western blot - 91 -](#_Toc504580862)

[2.6. Purification - 92 -](#_Toc504580863)

[2.6.1. Polyhistidine tag-Nickel purification - 92 -](#_Toc504580864)

[2.6.2. Biotin-avidin purification - 92 -](#_Toc504580865)

[2.6.3. Filtration - 93 -](#_Toc504580866)

[2.6.4. Centrifugation - 93 -](#_Toc504580867)

[2.6.5. Size exclusion chromatography - 93 -](#_Toc504580868)

[2.7. CYP3A4 activity assay - 94 -](#_Toc504580869)

[2.8. Characterisation - 94 -](#_Toc504580870)

[2.8.1. Dynamic light scattering - 94 -](#_Toc504580871)

[2.8.2. Flow cytometry - 95 -](#_Toc504580872)

[2.8.3. Fluorescence microscopy - 95 -](#_Toc504580873)

[2.8.4. Transmission electron microscopy - 96 -](#_Toc504580874)

[3. Cloning and Expression of CYP3A4 - 97 -](#_Toc504580875)

[3.1. Introduction - 99 -](#_Toc504580876)

[3.2. Aims - 100 -](#_Toc504580877)

[3.3. Experimentation - 101 -](#_Toc504580878)

[3.3.1. Cloning - 101 -](#_Toc504580879)

[3.3.2. Expression - 101 -](#_Toc504580880)

[3.4. Cloning - 102 -](#_Toc504580881)

[3.5. Expression - 102 -](#_Toc504580882)

[3.6. Discussion - 105 -](#_Toc504580883)

[3.7. Summary - 108 -](#_Toc504580884)

[4. Expression and Purification of CYP3A4 Microsomal Monooxygenase Proteopolymersomes - 109 -](#_Toc504580885)

[4.1. Introduction - 111 -](#_Toc504580886)

[4.2. Aims - 113 -](#_Toc504580887)

[4.3. Experimentation - 114 -](#_Toc504580888)

[4.3.1. Protein co-expression - 114 -](#_Toc504580889)

[4.3.2. Fluorescent vesicles - 114 -](#_Toc504580890)

[4.3.3. Centrifugation purification - 115 -](#_Toc504580891)

[4.3.4. Filtration purification - 115 -](#_Toc504580892)

[4.3.5. Biotin-avidin purification - 115 -](#_Toc504580893)

[4.3.6. SEC purification - 116 -](#_Toc504580894)

[4.4. Centrifugation purification - 116 -](#_Toc504580895)

[4.5. Filtration purification - 122 -](#_Toc504580896)

[4.6. Biotin-avidin purification - 123 -](#_Toc504580897)

[4.7. SEC purification - 125 -](#_Toc504580898)

[4.8. Discussion - 127 -](#_Toc504580899)

[4.9. Summary - 130 -](#_Toc504580900)

[5. Testing the Activity of the CYP3A4 Proteopolymersome - 132 -](#_Toc504580901)

[5.1. Introduction - 134 -](#_Toc504580902)

[5.2. Aims - 135 -](#_Toc504580903)

[5.3. Experimentation - 136 -](#_Toc504580904)

[5.3.1. Protein co-expression - 136 -](#_Toc504580905)

[5.3.2. Activity analysis - 137 -](#_Toc504580906)

[5.4. Total CF CYP3A4 monooxygenase activity - 137 -](#_Toc504580907)

[5.5. Altering conditions of expression - 140 -](#_Toc504580908)

[5.6. Centrifuged CF CYP3A4 monooxygenase activity - 140 -](#_Toc504580909)

[5.7. Filtered CF CYP3A4 monooxygenase activity - 145 -](#_Toc504580910)

[5.8. Avidin purified CYP3A4 proteopolymersome activity - 146 -](#_Toc504580911)

[5.9. Discussion - 151 -](#_Toc504580912)

[5.10. Summary - 154 -](#_Toc504580913)

[6. Characterisation of the CYP3A4 Proteopolymersome - 155 -](#_Toc504580914)

[6.1. Introduction - 157 -](#_Toc504580915)

[6.2. Aims - 159 -](#_Toc504580916)

[6.3. Experimentation - 159 -](#_Toc504580917)

[6.3.1. Avidin purification of proteopolymersomes - 159 -](#_Toc504580918)

[6.3.2. Stability analysis - 160 -](#_Toc504580919)

[6.3.3. Flow cytometry - 160 -](#_Toc504580920)

[6.3.4. DLS analysis - 160 -](#_Toc504580921)

[6.3.5. TEM Imaging - 161 -](#_Toc504580922)

[6.4. Stability of the proteopolymersome - 161 -](#_Toc504580923)

[6.5. Flow cytometry analysis - 167 -](#_Toc504580924)

[6.6. TEM Imaging - 173 -](#_Toc504580925)

[6.7. Discussion - 175 -](#_Toc504580926)

[6.8. Summary - 180 -](#_Toc504580927)

[7. Development of Peptide-Based Polymersome Anchors - 181 -](#_Toc504580928)

[7.1. Introduction - 183 -](#_Toc504580929)

[7.2. Aims - 186 -](#_Toc504580930)

[7.3. Experimentation - 187 -](#_Toc504580931)

[7.3.1. Cloning, expression and purification - 187 -](#_Toc504580932)

[7.3.2. Reconstitution and SEC - 187 -](#_Toc504580933)

[7.3.3. Fluorescence microscopy - 188 -](#_Toc504580934)

[7.4. Expression and purification - 188 -](#_Toc504580935)

[7.5. Reconstitution and SEC - 190 -](#_Toc504580936)

[7.6. Fluorescence microscopy - 195 -](#_Toc504580937)

[7.7. Co-translational insertion - 198 -](#_Toc504580938)

[7.8. Discussion - 199 -](#_Toc504580939)

[7.9. Summary - 203 -](#_Toc504580940)

[8. Discussion - 205 -](#_Toc504580941)

[8.1. Introduction - 207 -](#_Toc504580942)

[8.2. Cloning and expression of cytochrome P450 enzymes - 207 -](#_Toc504580943)

[8.3. Purification and activity of CYP3A4 proteopolymersomes - 211 -](#_Toc504580944)

[8.4. Characterisation of proteopolymersomes - 218 -](#_Toc504580945)

[8.5. Peptide anchors - 220 -](#_Toc504580946)

[8.6. Summary - 221 -](#_Toc504580947)

[9. Future Perspectives and Concluding Remarks - 223 -](#_Toc504580948)

[9.1. Introduction - 225 -](#_Toc504580949)

[9.2. Improving the CYP3A4 Proteopolymersome - 225 -](#_Toc504580950)

[9.3. Utilising the CYP3A4 Proteopolymersome - 227 -](#_Toc504580951)

[9.4. Conclusion - 230 -](#_Toc504580952)

[Publications - 231 -](#_Toc504580953)

[Appendix - 232 -](#_Toc504580954)

[References - 238 -](#_Toc504580955)

# Abstract

Microsomal monooxygenases are one of the most important subfamilies of cytochrome P450 enzymes. These membrane-associated enzymes, and their reductases, are of high research importance for the role they play in drug metabolism and interactions. One of the most active member of this family of enzymes is the cytochrome P450 3A4 enzyme, which metabolises over half of all current drugs available on the market. Being able to study and understand how this enzyme works is important for future drug development, including the creation of novel targets and predicting possible side effects. The creation of a functional CYP3A4 enzyme model with biological fidelity has been the goal of many research groups. However, challenges have been faced in the development of such models, both in creation of a stable and bio-compatible membrane environment for the anchoring of the enzymes, as well as having sufficient enzyme activity. Here we present a system created using cell-free expression system supplemented with polymer vesicles, so called polymersomes, for simple and quick co-expression of cytochrome P450 3A4, cytochrome P450 reductase and cytochrome *b5*. The availability of polymersome in the expression mixture allowed for the co-insertion of the enzymes into the polymersome membranes. The resultant proteopolymersomes were readily purified, and analysis showed them to be active and stable while maintaining the structural integrity of the polymersome throughout the entire process. This developed system has a high value in pharmaceutical research and is not only applicable to other members of the microsomal monooxygenase family of enzymes, but also to all other high interest membrane associated enzymes.

**Chapter 1:**

# Literature Review

1.1 SUMMARY

1.2 MICROSOMAL MONOOXYGENASES

1.3 CELL-FREE PROTEIN EXPRESSION

1.4 BIOMIMETIC CELL MEMBRANES

1.5 AIMS

List of Abbreviations

|  |  |
| --- | --- |
| AMV – Alfaalfa mosaic virus  C*b5*– Cytochrome *b5*  CF – Cell free  CHO – Chinese hamster ovary  CMC – Critical micelle concentration  CPR – Cytochrome P450 reductase  CrPV – Cricket paralysis virus  CYP3A4 – Cytochrome P450 3A4  EGFP – Eukaryotic green fluorescent protein  EMCV – Encephalomyocarditis virus  epPCR – Error prone polymerase chain reaction  FAD – Flavin adenine dinucleotide  FMN – Flavin mononucleotide  GPCR – G protein coupled receptor | IRES – Internal ribosome entry site  NADP – Nicotinamide adenine dinucleotide phosphate  PBD – Polybutadiene  PCL - Polycaprolactone  PDMS - Polydimethylsiloxane  PEG – Polyethelyne glycol  PMOXA – Poly 2-methyloxazoline  PEO – Polyethelyne oxide  PTM – Post translational modification  SDM – Site directed mutagenesis  TMV – Tobacco mosaic virus  uAA – Unnatural amino acid  UTR – Untranslated region  WT – Wild type |

List of Tables

**Table 1.1:** Classification of cytochrome P450 enzymes based on redox partners.

**Table 1.2:** Selected list of human CYP3A4 substrates, inducers, and inhibitors.

**Table 1.3:** Comparison of the advantages and disadvantages of different CF systems.

**Table 1.4:** Vector elements utilised for improvement of protein transcription and translation, and increase in yields.

**Table 1.5:** Commercially available CF protein expression systems and their key features.

**Table 1.6:** Examples of polymers utilised in the synthesis of polymersomes.

**Table 1.7:** Examples of proteopolymersome production.

List of Figures

Figure 1.1: Microsomal electron transport chain, electron movement from the NADPH to the microsomal monooxygenase is shown in red.

Figure 1.2: Amino acid sequence of CYP3A4, showing the N-terminus membrane spanning domain, and the large cellular domain.

Figure 1.3: CYP3A4 microsomal monooxygenase model highlighting the electron transport chain from CPR and C*b5*.

Figure 1.4: Scheme of cytochrome P450 enzyme activity utilising CPR and C*b5* for first and second electron donation respectively

Figure 1.5: Overview of cell-free protein expression system, composition and applications.

**Figure 1.6:** Methods for the generation of ATP driving cell-free transcription and translation

**Figure 1.7:** Arrangement of the hydrophilic and hydrophobic regions on polymers, and their arrangement on vesicle membranes.

## Summary

Membranes are the gatekeepers of the cell. They are the barriers that both encircles and contains the intracellular milieu while also preventing its mixing with the extracellular environment. While this is the primary function of membranes and one that is vital for the maintenance of cellular function, it’s not its only function. Membranes also have a key role as a scaffolding environment for a wide range of structures, ranging from proteins to lipids. Membrane-associated proteins need membranes for their stabilisation and solubilisation. These membrane-associated proteins are separated into widely varying family groups based on their structures and with equally varying functions. Membrane protein receptors, such as G-protein coupled receptors, are responsible for extracellular signal transduction across the membrane giving the intracellular environment the ability to respond according to its extracellular environment. Structural proteins, such as adhesins, help stabilise and attach cells to their neighbouring cells and environments. Transport proteins, as the name suggests facilitate the movement, passively or actively, of chemicals across the membrane barrier. The final group of membrane-associated proteins are the membrane enzymes. These enzymes are a part of a wide range of families and have diverse structures and functions, including kinases, hydrolases, transferases and oxidoreductases.

One group of membrane-associated oxidoreductases, the microsomal monooxygenases, are members of the cytochrome P450 enzyme superfamily and have attracted a lot of attention since their discovery. These enzymes, associated to intracellular endoplasmic reticulum membranes, have been shown to play a key role in many metabolic processes, including lipids and steroids synthesis. In addition to their diverse endogenous substrate specificity, they have also been shown to have a key role in the metabolism of a wide range of exogenous compounds including many commonly prescribed medicinal drugs. It is of vital importance to study and understand the structure and function of these enzymes not only to better understand them but also to understand current drug metabolism, adverse drug side-effects and to aid future drug design. One avenue for this research is the development of high fidelity enzyme models based on their native cellular arrangements. The development of accurate and true to life *In vitro* models of microsomal monooxygenases would offer a simple yet valuable tool in pharmaceutical research and understanding the relationship between drug and enzyme. With *in vitro* models, research can be simplified, controlled and offer greater convenience when compared to *in vivo* and *ex vivo* models.

To develop such an *in vitro* model, it is required to first express the enzymes, and to present them in an environment that is similar to its native environment. While it is possible to use native membranes for the stabilisation of enzymes, artificial cell membranes have become a leading source for biomimetic membrane environments, desired for their often-simpler preparation methods, tunability and higher levels of stability. One such novel biomimetic membrane are polymer vesicles, so called polymersomes. These vesicles can offer similar environments to those required by membranes for protein anchoring, stabilisation and folding. While the creation of the artificial membranes is simple enough, localising the enzymes on the membranes presents challenges of stability, orientation and interaction. One such method of tackling this problem is the utilisation of cell-free protein expression systems. These expression systems are desired for their openness and modifiability, including the supplementation of polymersome vesicles for co-translational insertion of membrane enzymes using the present expression machinery.

The hypothesis of this project is that cell-free protein expression systems will offer a simple and time-saving method for both the creation of protein-polymer complexes and their purification, a superior alternative to the standard cell-based expression methodology of separate stages of expression, purification and reconstitution. The polymersomes will act as a suitable biomimetic environment that can support the direct incorporation of functional membrane bound enzymes, with superior stability and functional life span than lipid based membranes.

## Microsomal monooxygenases

### Introduction

Cytochrome P450s form a large family of oxidizing haemoprotein enzymes found across all domains of life. These enzymes are responsible for many physiological processes including the metabolism of endogenous compounds, giving it a key role in steroid synthesis and fatty acid synthesis, among others. Importantly, cytochrome P450 enzymes are also widely involved in the metabolism of many exogenous compounds. In humans, enzymes such as CYP3A4 and CYP2C9, interact with a significant percentage of known drugs and are responsible for their bio-activation as well as their inactivation and clearance. Similarly, in other domains of life cytochrome P450 enzymes are involved in the metabolism of many exogenous xenobiotic compounds as well as many biosynthetic pathways. The wide functionality and availability across many species makes cytochrome P450 enzymes an attractive target for study and use in biotechnology and biopharmaceutical industries.

Microsomal monooxygenases are significant members of cytochrome P450 family of enzymes. This group, as the name suggests, are found *in vivo* on endoplasmic reticulum membranes, and when purified *in vitro* are located on microsomal membranes, along with the FAD and FMN utilizing cytochrome P450 reductase for the transfer of electrons and the incorporation of one molecule of oxygen into the substrate. In addition, the haemoprotein cytochrome *b5* can be used to boost the reducing activity. The main significance of this group arises from how prominent and active its members are against many endogenous compounds as well as, even more importantly, exogenous drug compounds. Its members, such as CYP3A4, CYP1A2, CYP2D6 and CYP2C9 interact with the majority of commercial drugs, affecting key drug characteristics including potency, elimination and side-effects. Their broad substrate specificity has really highlighted the importance of the study of these enzymes and understanding how interactions between members occur, how reactions are driven and what effects are observed with natural variations. The study of these enzymes offers many challenges from their subcellular localization on membranes to the creation of functional monooxygenase systems capable of translating accurate metabolism research. The structures of the cytochrome P450 enzymes have multiple key features; substrate binding domain, nucleotide binding domain, reductase enzyme binding regions and their singular membrane-spanning anchor region. Each of these features has been the focus of a large amount of research towards understanding and optimizing enzyme function. The single-pass trans-membrane domain, a characteristic of the P450 microsomal monooxygenases and the reductase enzymes are an obstacle in the study of these enzymes. Studies on the microsomal monooxygenases originally carried out utilising native membrane preparations, have recently been performed in biomimetic membrane environments aiming to mimic the native environments and obtain an accurate representation of the true nature of the enzyme interactions and activities of the microsomal monooxygenases.

In this review the structures, functions and importance of the microsomal monooxygenases will be discussed with special focus on the CYP3A4 monooxygenase and the partner proteins; CPR and C*b5*, as well as how engineering and modification of these enzymes and their environments has aided our understanding and study of them.

### Naming and classification

Cytochrome P450 enzymes are a large and varied superfamily of heme-containing enzymes, found across all kingdoms of life. They are however not found in all species, for example, they are absent from *E. coli*. As such, there are multiple classification systems for grouping cytochrome P450 enzymes based on their origin (*e.g.* eukaryotic/prokaryotic), their redox partners or their localization. Based on the redox partners, there are four classes to which cytochrome P450 enzymes belong; class I, II, III and IV (table 1.1). The grouping based on this classification however holds a weak correlation with the substrate specificity of the enzymes, with some of the microsomal monooxygenases for example capable of metabolising a wide range of both endogenous and exogenous compounds.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Table 1.1: **Classification of cytochrome P450 enzymes based on redox partners.** | | | | |
| Class | **Description** | **Origin** | **Examples** | **References** |
|  |  |  |  |  |
| I | 3 Enzyme systems; a reductase (FAD cofactor containing flavodoxin protein), iron-Sulphur ferredoxin protein and the cytochrome P450 enzyme | Mitochondrial (Eukaryotic) CYPs, bacterial CYPs | CYP11A1, CYP101 | (Schiffler and Bernhardt, 2003) |
|  |  |  |  |  |
| II | 2 Enzyme systems; a reductase (FAD & FMN cofactor containing protein) and the cytochrome P450 enzyme. | Microsomal (eukaryotic) CYPs | CYP3A4, CYP2D6 | (Davydov, 2011) |
|  |  |  |  |  |
| III | Single enzyme system; reductase domain fused to the catalytic domain of CYP. | Bacterial CYPs | BM3 (CYP102A), CYP505A1 | (Munro et al., 2002, Girvan et al., 2006) |
|  |  |  |  |  |
| IV | Single enzyme system; FMN-containing reductase with a ferredoxin-like centre linked to a P450 in a single polypeptide | Bacterial CYPs | P450RhF | (Roberts et al., 2002) |

### Microsomal monooxygenases

Microsomal monooxygenases, a subset of cytochrome P450 enzymes, are present *in vivo* on the endoplasmic reticulum membranes of eukaryotic cells. These enzymes catalyse a specific reaction, as shown here:

RH + Reduced flavoprotein + O2 ⇋ ROH + Oxidised flavoprotein + H2O

Members of this enzyme group share similar structural features; a membrane anchoring N-terminus sequence, a flavoprotein (the cytochrome P450 reductase) binding site, and as with all cytochrome P450 enzymes, a heme binding pocket. The functionality of these enzymes relies on the creation of a complete electron transport chain, the so called microsomal electron transport chain (Cederbaum, 2015) (figure 1.1). Briefly, electrons donated from NADPH are passed onto the flavin cofactors; FAD and FMN of the reductase. With close interaction between the reductase and microsomal monooxygenase, electrons can be transported to the cytochrome P450 enzyme to trigger the monooxygenation of the substrate.

|  |
| --- |
| Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-05-04 at 11.22.22.png |
| Figure 1.1: Microsomal electron transport chain. Electron movement through the CPR from the NADPH to the microsomal monooxygenase is shown in red. |

The vital importance of this group of enzymes is highlighted by their broad functionality in many physiological processes, including an essential role in steroid synthesis, including vitamin metabolism and lipid metabolism (Prosser and Jones, 2004, Catharine Ross and Zolfaghari, 2011, Pinto and Cooper, 2014, Bishop-Bailey et al., 2014). The importance of the role the microsomal monooxygenases play in the metabolism of a large number of diverse exogenous compounds for many organisms is highlighted by their wide substrate specificity and their prolific ability to metabolise many compounds. It is acknowledged that over 70% of current drugs are metabolic targets of microsomal monooxygenases, the most active enzymes are CYP 1A2, 2C9, 2C19, 2D6 and 3A4, with CYP3A4 being responsible for around half of all microsomal monooxygenase mediated metabolic clearance (Rendic and Guengerich, 2015).

### Cytochrome P450 3A4

CYP3A4 (EC 1.14.13.97) is one of the most studied members of the microsomal monooxygenases. It is the most abundantly expressed cytochrome P450 in the human liver, the major site of drug metabolism, highlighting the prominent role it plays. A lot of research is being conducted to understand the structure of the enzyme and elucidate all the roles it plays in the metabolism of endogenous and exogenous compounds and removal of toxicity. Other members of the CYP3A family include, CYP3A5; a polymorphism that is only present in about 10-20 % of the population, and CYP3A7; a variant with significant activity in foetuses, however most activity is lost by adulthood (Guengerich, 1999).

#### Structure

The structure of CYP3A4, is similar to that of other members of microsomal monooxygenases (figure 1.2). Its N-terminus (amino acid residues: 2-22) contains a single membrane spanning domain, which anchors the enzyme on the endoplasmic reticulum. The majority of the enzyme (amino acid residues: 22-503) forms the large cellular domain, containing the substrate binding pocket and the metal iron of the heme cofactor binding site (amino acid residue: 442). Microsomal monooxygenases in general are translated with a leader sequence that is recognised by cellular machinery for co-translational insertion into the endoplasmic reticulum membrane, followed by a stop signal to prevent the passage of the whole peptide across the membrane, leaving the large catalytic domain of the enzyme in the cytosol environment (Sakaguchi et al., 1984, Sakaguchi et al., 1987). Microsomal monooxygenases also have additional membrane interacting hydrophobic peptide sequences that can bind to the membrane surfaces (Pernecky et al., 1993, Williams et al., 2000).

#### Function

CYP3A4 is expressed by mammalian species, localized within the liver, gut (intestinal epithelium, bile duct, gastric epithelium), prostate and brain (Lacroix et al., 1997, Thummel, 2007, Dutheil et al., 2008). Its expression is induced by glucocorticoids and some drug targets such as rifampicin and phenytoin (Guengerich, 1999, Zhou, 2008). Its catalytic substrate specificity is very broad, even when compared to most cytochrome P450 enzymes. Its range of substrates ranges from the small (~150 Mr *e.g.* acetaminophen) to the larger compounds (~1200 Mr *e.g.* cyclosporin A). A small selection of CYP3A4 substrates, and expression inducers and inhibitors have been presented in table 1.2. The cellular metabolic pathways of CYP3A4 include C- and N-dealkylation, C-hydroxylation, dehalogenation, and nitroreduction (Li et al., 1995).

|  |
| --- |
| **Endoplasmic reticulum**  **Variants** |
| Figure 1.2: Amino acid sequence of CYP3A4, showing the N-terminus membrane spanning domain, and the large cellular domain. |

Current *in vitro* strategies for the study of CYP3A4 metabolism of substrates usually involve the usage of microsomal preparations from human hepatocytes, as well as expression in genetically engineered cell lines (human, yeast and insect cells) (Li et al., 1995). Multiple microsomal preparation protocols are available, and are usually tailored for specific research purposes. Generally, these protocols involve a homogenization step for the break-up of cells in a suitable buffer, followed by serial centrifugation steps for purification of the microsomes from the cellular debris and storage at a suitable temperature.

|  |  |  |  |
| --- | --- | --- | --- |
| Table 1.2: Selected list of human CYP3A4 substrates, and expression inducers and inhibitors. | | | |
| Substrates | Expression | | |
| Drugs | Inducers | | Inhibitors |
| Anaesthesia *e.g*. codeine, lidocaine, methadone  Antibiotics *e.g.* erythromycin, quinine  Antihistamines *e.g*. astermizole, chlorpheniramine  Cardiovascular *e.g*. nifedipine, quinidine  Chemotherapy *e.g.* tamoxifen, doxorubicin  Steroids *e.g.* hydrocortisone, testosterone | Barbiturates  Carbamazepine  Glucocorticoids  Modafinil  Phenobarbital  Phenytoin | Clarithromycin  Itraconazole  Ketoconazole  Nefazodone  Ritonavir  Saquinavir | |
| Carcinogens & pollutants |
| Benzopyrene |
| Pesticides |
| Parathion, Aldrin |

The isolated preparations have shown that to be functional, it must exist as part of a multienzyme system composed of multiples of microsomal monooxygenases, reductases and cytochrome *b5*, and in varying ratios, usually with the monooxygenases in excess (Davydov, 2011). It’s also been shown that microsomal monooxygenases exist in a homomeric and heteromeric state natively with multiple different monooxygenases cooperating for metabolism of substrates.

### Cytochrome P450 reductase

CPR is a multi-enzyme reductase, supplying the electrons required for catalytic activity to all microsomal monooxygenases. Similar to the microsomal monooxygenases, it is anchored to the endoplasmic reticulum membrane with a single transmembrane spanning region located within the N-terminus (amino acid residues: 22-42), with a short preceding luminal peptide (amino acid residues: 2-21), and a much larger cytoplasmic C-terminus region (amino acid residues: 43-677). It contains three cofactor binding pockets; NADPH (amino acid residues: 596-606), FMN (amino acid residues: 86-91, 138-141 and 173-182), and FAD (amino acid residues: 454-457, 472-474 and 488-491). Its role is central within the electron transport chain, transferring the electrons from NADPH cofactors, through FAD and FMN, to the microsomal monooxygenases (as shown in figure 1.1).

#### Interactions with CYP3A4

CYP3A4 metabolic activity, both against endogenous and exogenous xenobiotics, is highly dependent on CPR and its transfer of electrons for catalytic function. Its localisation on the endoplasmic reticulum membrane, the same localisation as all microsomal monooxygenases, allows it to have a tight protein-protein interaction with CYP3A4, as well as the other microsomal monooxygenases (figure 1.3). The arrangement of the *in vivo* monooxygenase system usually has the cytochrome P450s in excess, from 10-fold to 30-fold, compared to the CPR. However, in *in vitro* models, it has been shown higher levels of monooxygenase activity when CPR is present in excess (Lee and Goldstein, 2012).

### Cytochrome *b5*

C*b5* is a heme-containing, electron donor to the microsomal monooxygenase enzymes. It has two variants; one membrane associated to the endoplasmic reticulum, while the other is found within the cytoplasm. The membrane-associated variant is the one of interest in the role it plays in electron donation and its interactions with microsomal monooxygenases, especially towards CYP3A4. C*b5* type A, the microsomal membrane bound variant, is a 134 amino acid protein with a single C-terminus membrane anchoring region (amino acid residues: 109-131). This is in contrast to microsomal monooxygenases, CPR and C*b5* reductase which all have their membrane spanning region at the N-terminus. It is thought that C*b5* has its membrane binding domain forming a hairpin loop with the C-terminus end existing in the cytoplasmic side. Its N-terminus is dominated by the heme binding domain (amino acid residues: 9-85).

|  |
| --- |
| Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-04-13 at 10.06.51.png |
| Figure 1.3: CYP3A4 microsomal monooxygenase model highlighting the electron transport chain from CPR and C*b5*. |

Its role in the electron transport is that of an in-between connecting reductases and monooxygenases, as it can accept electrons from both CPR and C*b5* reductases, before passing it forward onto the monooxygenases (Vergeres and Waskell, 1995, Schenkman and Jansson, 2003) (figure 1.3). It additionally plays a significant role in the metabolism of many compounds, including the biosynthesis of lipids, cholesterol and glycoproteins, as well as in the reduction of cytochrome P450 enzymes.

The mechanisms of action of C*b5* with cytochrome P450 enzymes are heavily studied and thought to be a combination of electron transfer and allosteric stimulation. C*b5* primary electron role is thought to be on transferring the second electron to cytochrome P450 enzymes (figure 1.4). C*b5* is a single electron acceptor, converting itself from a ferric (Fe3+) haemoprotein to a ferrous (Fe2+) haemoprotein. With cytochrome P450 enzymes the ferrous C*b5* is capable of utilising this electron and transferring it to the oxyferric (oxygen bound, Fe3+) cytochrome P450 complex, a usually rate-limiting step for microsomal monooxygenases (Schenkman and Jansson, 2003). It is thought the transfer of the second electron from C*b5* stabilises the cytochrome P450 complex, preventing the release (uncoupling) of the superoxide anion from the cytochrome P450 complex (Perret and Pompon, 1998). This enhancement of coupling directly leads to higher levels of product release. Some aspects of C*b5* mediated cytochrome P450 activity are still not clear. It is known that NADH mediated reduction of C*b5* by C*b5* reductases is faster than NADPH reduction via CPR, yet however the enhancing effect of C*b5* on microsomal monooxygenases is minimally observed when it is carried out via NADH reduction (Perret and Pompon, 1998, Cederbaum, 2015).

|  |
| --- |
|  |
| Figure 1.4: Scheme of cytochrome P450 microsomal monooxygenase enzyme activity utilising CPR and C*b5* for first and second electron donation respectively |

#### Interactions with CYP3A4

C*b5* is often regarded as an enhancer to CYP3A4 activity, however all the avenues by which this happens have not yet been clarified and is usually dependent on the isoform of monooxygenase. It has been shown that C*b5* interactions with monooxygenases, such as CYP3A4, exerts an allosteric effect altering the structure for a more optimal conformation. Studies have shown that heme deficient C*b5* (apo-C*b5*) can have similar effects to those containing the heme cofactor, including boosting metabolic activity of CYP3A4 and other microsomal monooxygenases (Yamazaki et al., 2002, Locuson et al., 2007). It has been suggested that the apo-C*b5* possibly exert their enhancing effects by accepting heme groups from the microsomal monooxygenases and therefore allowing electron transfer (Guryev et al., 2001). To test this theory, addition of heme scavenging apo-myoglobin was carried out and showed the removal of the enhancing properties of apo-C*b5*. However, by analysis of whether such a heme transfer was possible with CYP3A4, it was shown that only a small percentage of heme groups were capable of being transferred from the monooxygenase to the C*b5*, and indeed it showed that the boosting effect of apo-C*b5* was not seen with all CYP3A4 metabolites (Yamazaki et al., 2001, Yamazaki et al., 2002).

The enhancing effect of C*b5* on CYP3A4 was observed on the rate of testosterone hydroxylation by recombinant CYP3A4 and purified human liver microsomes (Yamazaki et al., 1996). Results showed an increase in activity with the use of CPR over C*b5* reductase, and further enhancement was observed when NADPH cofactor was used rather than NADH. Removal of C*b5* from the system lead to a 90% decrease in observed activity, and no activity was observed when CPR was also removed. These results indicate the important role C*b5* plays in boosting monooxygenase activity and the importance of electron flow through it from CPR to the monooxygenase.

### Enzyme engineering

Engineering of enzymes for improved or novel activity and functions has been widely carried out on a large variety of enzymes across all families with broad success, the cytochrome P450s are no different as well. For large scale applications in research and industry, the optimization and modification of enzyme characteristics will always be desired. The techniques of random evolution and rational, targeted mutations are the answer to the problems faced when engineering cytochrome P450s. These techniques have been used to improve enzymatic activity and stability, and even adding metabolic activity towards new substrates. Rational design of enzymes utilizes the current knowledge on enzyme structure and function to theorize on the roles of peptides in the way the enzyme works, and how its alteration can affect its function for the better. Directed evolution is a broader approach to engineering, relying on the mutation of the whole or part of the enzyme while monitoring the changes in the characteristics of the enzymes, usually with the enzyme phenotype remaining linked to the genotype, so those enzymes with improved characteristics can be selected, isolated and analysed (Tee and Wong, 2013).

#### Activity and structure engineering

One of the main targets for engineering is the improvement of the catalytic activity of enzymes towards specific substrates. The microsomal monooxygenase CYP2A6 has undergone random mutagenesis using epPCR to improve catalytic activity towards one of its substrate, coumarin (Lee et al., 2014). Multiple mutations affecting the substrate binding site were found to improve the coumarin hydroxylation turnover rate. Interestingly, it is not always required to mutate the active site of an enzyme for the improvement of its activity. The role of residues in redox partner interactions and substrate recognition and binding are all important, and must be considered for optimal enzyme activity. This was seen with epPCR of CYP2B1 identifying multiple mutation points away from the active site that showed an improved activity for multiple substrates, including a quadruple mutant with a 6-fold higher catalytic activity in hydrogen peroxide dependent oxidation (Kumar et al., 2005). Similarly, CYP3A4 underwent epPCR treatment aimed at improving peroxide-supported substrate oxidation, coupled with targeted SDM (Kumar et al., 2006), based on the knowledge of point mutations in CYP2D6 that improved catalytic activity. The combined epPCR and SDM mutants showed up to 11-fold improvement in catalytic efficiency of some mutants. More importantly is the fact that the mutants are capable of supporting peroxide mediated oxidation more efficiently than WT CYP3A4, alleviating the need for the typical CPR, C*b5*and the expensive NADPH cofactor. In CYP1A2, random mutagenesis and screening generated multiple mutants that possessed higher activity (Kim and Guengerich, 2004). Analysis of the mutants and modelling based on the structure of another enzyme, CYP2C5, suggested the mutations to be unrelated to the substrate binding site. One of key takeaways from these studies is the importance of utilizing all avenues available for understanding the relationship between structure and function and exploiting it for gain, whether as seen with CYP3A4 for suggestions of advantageous point mutations, or with CYP1A2 for helping to understand the enzyme structure and role of the mutations in its function.

#### N-terminus modifications

The membrane spanning regions of the microsomal monooxygenases, CPR and C*b5* (at the C-terminus) have attracted special interest due to its importance in membrane binding and the anchoring of enzymes to the endoplasmic reticulum. Varying modifications and engineering of these membrane spanning regions have been carried out to study their effects on expression, activity and protein-protein interactions. With microsomal monooxygenases, the N-terminus has been altered and deleted to study its importance in enzyme aggregation, activity and membrane anchoring. One study with wild-type CYP3A4 and four altered mutants with altered membrane spanning domains found one mutant, the CYP3A4 Nf14, with 9 amino acid deletions, was capable of higher levels of recombinant expression in *E. coli* compared to the WT, and with activity levels comparable to those seen in human liver microsomes (Gillam et al., 1993). The N-terminus of CYP3A4 Nf14, which is the amino acid sequence MALLLAVFL, was first obtained from bovine CYP17A1 through substitution of the second amino acid residue from a tryptophan to an alanine, and *E. coli* codon optimisation of the following 5 residues (Barnes et al., 1991). Subsequently, alignment of the N-terminus sequence from the altered bovine CYP17A1 was applied to human microsomal monooxygenases. In CYP1A2 it has increased expression levels significantly and maintained activity (Fisher et al., 1992, Sandhu et al., 1994). Similar modification of the N-terminus was carried out on CYP3A5 showing possible optimisation of CYP3A family members’ reconstitution through sequence modification (Gillam et al., 1995). With CYP2E1, removal of 29 amino acids from the transmembrane spanning region and signal peptide has been shown to be functional, indicating that the N-terminus plays little role in the catalytic activity of the enzyme, however about two-thirds of the enzyme was still found on the membrane fraction rather than in the cytoplasm (Larson et al., 1991, Pernecky et al., 1993). Deleterious effects have also been observed in modification of the N-terminus of enzymes. Removal of the post transmembrane spanning domain amino acid sequence PPGP, which is found conserved in CYP1 and CYP2 enzymes, has shown to maintain the enzyme’s localisation on the endoplasmic reticulum membrane, however as a result of the alteration the protein is less stable and the catalytic activity has been abolished (Szczesna-Skorupa et al., 1993). Other avenues have since been developed for optimisation of cytochrome P450 expression, especially in poorly equipped protein expression system such as *E. coli*, including the supplementation of expression media with hemin and heme precursors, use of chaperones, and recently the utilisation of non-native membranes as stabilising environments for the enzymes.

The N-terminus of microsomal monooxygenases has been shown to affect its aggregation abilities and hinder its ability to form heterooligomer enzyme complexes. Mixtures of CYP2C9 and CYP3A4 in presence of low detergent concentrations have been shown to produce mixed aggregate complexes, while truncation of the N-terminus domains had limited oligomerization (Subramanian et al., 2010). Removal of the N-terminus, while it does aid in altering the localisation of the microsomal monooxygenase to the cytoplasm, it does not abolish ability of microsomal monooxygenases to bind to membranes. The four constructs of CYP3A4 with varying levels of N-terminus truncations and alteration were all found to be present on the membranes (Gillam et al., 1993). Other sequences within the monooxygenases have been reported to play a role in protein-membrane interaction and association (Davydov, 2011). Similarly, studies on CYP2B4 and CYP2E1 found it likely that other hydrophobic membrane interacting regions exist (Pernecky et al., 1995). Focus on finding these regions have pointed towards hydrophobic interactions of the F/G loop (Cosme and Johnson, 2002, Williams et al., 2000), as well as possible electrostatic interactions (von Wachenfeldt et al., 1997, Davydov, 2011).

The goal of solubilising the cytochrome P450 enzymes has opened the way for their crystallisation and the ability to determine their structure and understand how they functions, usually a difficult task for such lipophilic, aggregating membrane proteins. Crystallisation of the microsomal monooxygenases has shown their ability to undergo conformational changes in solution. In CYP2B4, it was shown that monooxygenases have open and closed conformations (Scott et al., 2003), while in CYP3A4, the conformation alteration was dependent on the bound substrate, while some substrates such as progesterone reported minor conformational changes, higher levels of flexibility and major conformational changes were seen with the substrates ketoconazole and erythromycin (Williams et al., 2004, Ekroos and Sjogren, 2006).

#### CYP-Reductase fusions

Looking at cytochrome P450s in nature, BM-3 has always been an attractive target for research, and especially for industrial applications. The fusion of the reductase domain to the catalytic domain on a single peptide is a highly valuable characteristic, and it is this “self-sufficiency” that simplifies the use of the enzyme in research, and with little surprise there is a multitude of current biotechnological and industrial applications (Chefson and Auclair, 2006), as well as ongoing research into future applications. There is a push as well to replicate the self-sufficiency of BM-3 for microsomal monooxygenases. Multiple attempts have been made to fuse the cytochrome P450s to their redox partners, with the aim of introducing novel applications in the fields of medicine, biosynthesis and bioremediation.

Early attempts at this include the fusion of the P450 enzyme CYP2D6 to the human CPR (Deeni et al., 2001). For this fusion, the N-terminus of the cytochrome P450 was joined to the C-terminus of the CPR and expressed in *E. coli,* showing similar activity towards metabolites as both enzymes co-expressed. Similarly, the membrane spanning C-terminus of C*b5* which interacts has been fused to CPR, replacing the membrane-spanning N-terminus of the reductase (Gilep et al., 2001). The fusion flavoprotein had varying activity when coupled to different cytochrome P450s; CYP4A1 and CYP17A showed higher levels of activity, but less with CYP3A4, highlighting the importance of the surface structure for protein-protein interactions and electron transfer. These fusion hybrid although novel and offer great potential, they are not yet capable of competing with native fusion cytochrome P450s like BM-3 (Gillam, 2007).

### Future directions and research

The central role microsomal monooxygenases play in human drug metabolism is very clear and has in turn made these enzymes a major subject for studies of drug design, adverse drug effects and protein-drug interactions. The membrane-association nature of microsomal P450s and the high tendency to form aggregates in solution hinder the mechanistic characterisation of this complex system. Many questions with regard to microsomal monooxygenases still need to be answered; including how do microsomal cytochrome P450 form homo- or hetero-oligomeric enzymes, and what role do these multi-enzyme complexes play in drug metabolism, and their interactions with reductases and cellular regulatory elements. As such there has been a push by many groups to utilise stable and robust biomimetic membrane elements, including liposomes, nanodiscs and polymersomes, with the goal of offering a stabilising environment outside the confines of the cell. Some research groups have also begun to look at utilisation of cell-free protein expression systems for the expression of microsomal cytochrome P450s in an open, controllable and modifiable environment for targeted research.

## Cell-free protein expression

### **Introduction**

There are multiple biological systems utilised for the expression of proteins. Commonly utilised cellular systems include the bacterial workhorse *Escherichia coli* or the eukaryotic *Saccharomyces cerevisiae*, Chinese hamster ovary and insect cells. Nowadays protein expression is a staple of most molecular biology labs, and there have been numerous advances in the field of protein expression (*e.g.* media, vectors and expression strain optimisations), but in general, the principles of expression have remained the same. The gene of the protein of interest is cloned into a suitable vector containing multiple cloning sites with an upstream promoter for control and drive of expression levels, a selection marker commonly an antibiotic resistance gene and finally an origin of replication. The designed vector is inserted into a suitable expression host. The protein expression host used would depend on factors such as the protein origin, its structure and the required PTMs. Protein expression has been a valuable resource for a huge number of scientific fields; recombinant protein expression in commonly used expression strains such as *E. coli* and CHO have paved the way for large scale protein expression for many bio-applications including biopharmaceuticals and industrial bio-catalysis. Protein expression has been utilised as well for protein characterisation and structural analysis which has elucidated many proteins’ functions and role in physiology therefore aiding in drug design and development.

From those cell-based protein expression platforms a new type of expression system was developed, one where there would be utilisation of the cellular extracts, but without the enveloping and constricting cellular membranes. These systems were called the CF expression systems. CF systems uniquely offer an expression platform that can be tightly controlled and altered accordingly, something that is a lot more difficult to offer in cellular expression. This difference is a key advantage that is well exploited; currently there are many CF expression systems being utilised for their ability to express unnatural amino acids, difficult and toxic proteins, as a tool for synthetic biology and in the development of minimal synthetic cells. The success of the wide-scale applications seen with CF systems has been through the effort of highlighting each CF expression system component such as lysates extract, sources, vectors and post-translational machinery, and optimising it to create highly efficient protein expression systems. In this literature review, CF systems from different sources will be examined for their benefits and applications as well as looking at future avenues for optimisation of CF expression systems utilisation.

### **Cell-**free **expression**

The cellular model for protein expression is well characterised and utilised. Alternatively, CF expression is still relatively new and under-utilised (figure 1.5). The dissimilarity of CF expression from cell-based protein expression is the key aspect of its attraction. The lack of cellular membranes enveloping the protein expression platform opens up CF expression to a multitude of new research opportunities and novel applications. Another important difference of CF systems is the time scale of the expression. In comparison with other commonly used protein expression systems, CF expression can have the shortest timescale. It can be possible to go from DNA to expressed protein within the range of a few hours, a significantly shorter timescale than that of bacterial (days) and eukaryotic (several days-week). The volume of expression mixture and the amount of protein produced should be carefully considered. Generally, CF expression is on a smaller scale than that of cell-based expression, although there is potential for scale-up but with added complexity, time and costs.

|  |
| --- |
| A close up of a map  Description generated with high confidence |
| Figure 1.5: Overview of cell-free protein expression system, composition and applications. |

Broadly, CF expression systems can be separated into two main categories; linked or coupled. Linked CF expression has separate environments for mRNA transcription and protein translation. The generation of mRNA is carried out independently, in cell-based systems or through the use of RNA polymerases, possibly followed by purification, ready for use with CF systems for protein translation. In coupled systems, the DNA is added to the system and the transcription and translation occur concurrently in the same environment.

CF protein expression characteristics such as lysate extract, vector, and energy sources have to be carefully considered when planning CF expression. The main characteristic is the composition of the CF expression mixture, *i.e.* the cellular lysates that form the platform for protein expression. CF mixtures are composed of cellular lysates or extracts originating from some of the commonly used cellular expression systems (*e.g. E. coli* and *Spodoptera frugiperda*). To ensure cellular lysates are functional and active in protein production, they must contain all the necessary components required to be able to transcribe and translate DNA to mRNA to protein efficiently including nucleotides and amino acids, to have all the necessary machinery such as RNA polymerases and ribosomes and the energy to drive the process in the form of ATP (figure 1.5). As well as being functional, cellular lysates also have to be suitable. The suitability of a CF system in expressing proteins is an important consideration due to the varying origins of target proteins and how different biological systems utilise cellular machinery to modify and process these proteins, and how capable they are of coping with the pressures of their expression. It is therefore a very important aspect of CF protein expression to have a reaction mixture possessing all required components to be functional and suitable.

### Extracts and lysates

There are a wide range of possible expression strains that can be used in the design and creation of CF expression systems, however when weighing their advantages and disadvantages, four extracts are more commonly utilised; the prokaryotic *E. coli* extract, and the eukaryotic wheat-germ extract, rabbit reticulocyte lysates and insect cell extract (*S. frugiperda* Sf21) (Carlson et al., 2012). Each of these CF systems has their own advantages and disadvantages (table 1.3). The prokaryotic *E. coli*, a workhorse for protein expression shares similar advantages and disadvantages for CF expression as it has for cell-based expression. *E. coli* extracts can be easily prepared at low costs for high level of protein production, have some well-known and modifiable biochemical and protein production pathways and are capable of folding complex proteins. However, its disadvantages are highlighted with the expression of complex proteins as it lacks post-translational machinery required for many eukaryotic proteins activity. The eukaryotic systems alternatively are much more capable of producing complex proteins with less aggregation and more varied PTMs (Gagoski et al., 2016, Brodel et al., 2013), however their disadvantages include the high costs and complex modifications required to produce a CF system with high levels of protein production.

#### *E. coli* extract

The use of *E. coli* cell extracts for cell free protein production highly popular and has been so for many decades. There are many aspects to the extracts which have been the focus of research groups, including energy source, methods of extract preparation and the vectors, promoters and polymerases utilized.

|  |  |  |
| --- | --- | --- |
| **Table 1.3:** Comparison of the advantages and disadvantages of different CF systems. | | |
| **Extract / lysates** | **Advantages** | **Disadvantages** |
| ***E. coli*** | High yields  Wealth of genetic and strain information  Simple and quick preparation  Relatively cheap cost of system | No PTMs  Limited by prokaryotic capabilities  Protein aggregation/misfolding |
| **Insect cells** | Able to carry out most PTMs  Process signal sequences  Express and fold complex proteins | Costly and time-consuming preparation  Non-mammalian  Limited genetic modification tools |
| **Wheat germ** | High yields  Express and fold complex proteins | No PTMs  Non-mammalian  Limited genetic modification tools |
| **Rabbit reticulocytes** | Native mammalian system  Minor PTMs | Requires additional components for PTMs (*e.g.* microsomes)  Limited genetic modification tools  Lower translation rates |
| **Human (HeLa/hybridoma)** | Native environment for expression of human proteins  Able to carry out complex PTMs  Less protein aggregation | Complex preparation protocols  Lower levels of expression |

Since the development of CF systems many protocols have been created and utilised with varying successes. One of the most commonly used and successful *E. coli* extract is the S30 extract. This cell extract system contains key features that aid its success, including reduced RNAase activity and deficiency in OmpT and lon proteases activity. Kigawa et al. (2004) have published a protocol for the preparation of *E. coli* S30 extract utilized by them for high throughput screening of proteins. The CF protein production system developed in that study is capable of producing human proteins that are correctly folded, as shown by NMR analysis. Since then there has been multiple reports published on improving S30 extract preparations including high temperature cultivation (Yamane et al., 2005) and streamlining and condensing the preparation steps (Liu et al., 2005). Different methods for the preparation of other *E. coli* CF extracts and improving the protein expression levels have been developed, an *E. coli* based CF system was created using a BL21 (DE3) strain (Michel and Wuthrich, 2012). This developed system was capable of producing purified folded proteins. Shrestha et al. (2012) focused on the cellular lysis methods in extract preparation, aiming for the development of a cheap and simple option with the focus on lysis techniques of sonication, bead vortex mixing, freeze-thaw and lysozyme. The best results were obtained through the use of sonication, with levels of recombinant EGFP produced comparable to other CF protein production systems. Recently it has been shown with scientific advances it is possible to produce a functional S30 extract from cell mass in under an hour (Krinsky et al., 2016).

An alternative to the S30 extract preparation from cellular lysates is the PURE (Protein synthesis Using Recombinant Elements) developed by Shimizu et al. (2001). The PURE system uses a reconstituted system of all the translation machinery of *E. coli* including translation initiation, elongation and termination factors, tRNA synthetases, T7 RNA polymerases and ribosomes. These components were expressed in *E. coli* cells, purified using affinity chromatography and reconstructed as a CF system. The specific construction of the PURE system allows for limiting the presence of expression inhibitors such as nucleases, proteases and nucleoside triphosphate hydrolysing enzymes, as well as suppressing the amber codon and the introduction of unnatural amino acids. Currently the PURE technology is being commercialised by biotech companies such as the New England Biolabs’ PURExpress *in vitro* protein expression system.

#### Insect cell extract

The advantages of using the eukaryotic insect cells for protein expression are well documented and has driven interest in developing similar insect cell lysate-based CF expression systems. One of the most successful system developed is the Sf21 (*Spodoptera frugiperda* 21) insect CF system developed by Ezure et al. (2006). Based on the Sf21 insect CF protein translation and glycosylation system (Tarui et al., 2000, Tarui et al., 2001), a simple and easy protocol was developed. The simple extract preparation is carried out using freeze-thaw cycle in a glycerol-based buffer. Optimisation of the translation system was implemented on the reaction components, and with the levels of luciferase mRNA being used as measure, the optimal concentrations of ATP, GTP, creatine phosphate, creatine kinase as well as ions including magnesium and phosphate were obtained. Previously it has been shown that the baculovirus expression system in insect cells has produced high levels of proteins *in vivo* under the action of a strong polyhedrin promoter (Smith et al., 1983). This was considered and utilised in the development of the Sf21 insect CF system, with the effects of multiple polyhedrin sequences on levels of expression examined (Ezure et al., 2006) as well as the development of pTD1 vector with a 5’ polyhedrin untranslated region (UTR) working as a translational enhancer (Suzuki et al., 2006). The enhancer sequences were successful in increasing expression levels not only in insect CF systems but also in wheat germ extracts and rabbit reticulocyte lysates systems. The success of this system has been widely exploited with protocols for use of the system published (Ezure et al., 2014). Commercially prepared Sf21 insect CF systems are also now available from Promega along with specialised vectors with 5’ and 3’ polyhedrin UTR regions for optimal protein expression.

Although not yet as frequent as those of the more commonly used *E. coli* CF systems, more insect lysate preparation protocols are being published. Protocols for preparation optimisation of insect cell extracts and its activity have been published (Tarui et al., 2001, Katzen and Kudlicki, 2006, Masuda et al., 2007, Richter et al., 2014). The ability of Insect cells to produce complex, correctly folded mammalian proteins with the appropriate post-translational modifications has driven research into the characterization and improvement of this CF system in the past and will continue to do so.

#### Wheat germ extract

Wheat germ CF systems have been widely in use since their development in 1964 (Marcus and Feeley, 1964) as one of the best and most efficient CF protein translation systems. Early usages of wheat germ CF systems were limited by ribosome inactivating proteins. Called tritin in wheat germ systems, these proteins were found to inhibit proteins by depurinating rRNA and inactivating the ribosomes (Coleman and Roberts, 1981, Takai and Endo, 2010). These proteins led to instability of the CF system and low protein expression. Solutions of this problems included attempting a continuous flow CF system with renewing of exhausted constituents coupled with removal of inhibitory by-products (Endo et al., 1992) which had already shown positive results in *E. coli* CF systems (Spirin et al., 1988, Kigawa and Yokoyama, 1991). Although continuous flow improved rate of protein translation, the best solution was the isolation of all inhibitory components from the wheat germ by washing before homogenization of the cells (Madin et al., 2000). The main target of the clearance was the endosperm that is the origin of tritin as well as other inhibitory elements. Through this protocol protein expression was observed for 4 hours in a batch system and for over 60 hours in a continuous system. Since then the wheat germ CF system has been commonly used for a wide range of applications, and as a eukaryotic expression system it is capable of producing complex, bioactive proteins (Goshima et al., 2008) and protocols for the preparation and utilisation of this system have since been published (Takai and Endo, 2010).

#### Rabbit reticulocyte lysate

Like wheat germ CF systems, rabbit reticulocyte lysates systems have been utilised for a few decades. Since being first developed by Pelham and Jackson (1976), evolution of the system has been carried out to improve performance. Nucleases, which catalyses the digestion of nucleic acid strands, is used to target and digest the endogenous mRNA of the rabbit reticulocytes, reducing the native protein translation to minimal levels. Although primarily pancreatic nucleases were utilised early on, micrococcal are more preferred now for their dependency on calcium ions, which allow for the inactivation of the nucleases through the use of calcium chelating agents. This nuclease treatment of lysates leads to minor levels of endogenous protein translation while having a minimal effect on exogenous protein translation (Pelham and Jackson, 1976). Other treatments have also been found to improve expression levels in rabbit reticulocyte lysates including the addition of the polyamine analogues including di-, tri-, tetra- and pentaamines (Snyder and Edwards, 1991). These analogues lowered the necessary concentration of the required magnesium ions during protein synthesis.

For optimal protein expression, factors including spermidine (Igarashi et al., 1980), GTP, glucose phosphate and protease inhibitors have been added to stimulate expression and minimise protein degradation, however expression levels of this system tend to still be lower than that achieved with the wheat germ extracts (Spirin, 2004, Katzen et al., 2005). Nonetheless rabbit reticulocytes lysates are well utilised as a CF system due to its ability to carry out some PTMs, a feature absent in wheat germ CF systems. They have been shown to carry out a wide range of PTMs, especially in the presence of canine microsomal membranes (Walter and Blobel, 1983, Gibbs et al., 1985, Sanford et al., 1991, Bulleid et al., 1992).

#### Other lysates

Outside the main four CF systems, other extracts have been developed for targeted functions. Human CF protein expression systems were developed for the expression of human proteins when it is important to gain accurate protein synthesis and post-translational modifications as seen with proteins expressed in human cells *in vivo*, as well as have true activity of the proteins*.* Multiple differing human cell lines have been utilised. The commonly used HeLa cancer cell line was used to develop a CF system with high levels of expression. HeLa CF protein expression system have been used to study many viruses, including the elucidation of replication and propagation methods of the encephalomyocarditis virus (EMCV), with viral particles successfully expressed using both RNA and DNA (Kobayashi et al., 2011, Kobayashi et al., 2012) and the synthesis of poliovirus to aid in the study of its method of infection and replication (Molla et al., 1991, Franco et al., 2005). Another source of human cells well utilised for CF protein expression originate from hybridoma cell line. Hybridoma CF systems are highly utilised for their ability to carry out complex glycosylation present widely across many eukaryotic proteins, and at a much more efficient rate than HeLa systems (Mikami et al., 2006a).

CHO cells are a highly utilised *in vivo* system for the production of therapeutic proteins in the biotechnology industry, and therefore have been a main target for the development of protein CF expression system. Progress on the development of a CHO *in vitro* production system has been slow; however major developments have been made in the creation of a high level functioning expression system. Brodel et al. (2014) have developed a highly specialised CHO CF system using CHO cell lysate coupled with specialised design expression vector and optimised ion, ATP and GTP concentrations all combined to create a CF system of protein synthesis levels comparable to those achieved with other eukaryotic CF systems, as well as the production of glycosylated proteins. The development of yeast based CF system has been carried out by Hodgman and Jewett (2013). The yeast species *S. cerevisiae* is a commonly used recombinant protein production strain, and there has been numerous research done exploiting its applications for the creation of an optimised and efficient CF system, including optimising extract preparation methods, vector design (Gan and Jewett, 2014, Hodgman and Jewett, 2014) and fermentation (Choudhury et al., 2014). Similarly, recent work has been carried out on the development of a novel CF protein expression system of plant origin, specifically from the tobacco Bright Yellow 2 cells, aimed at exploiting the plant’s simple and cost-effective fermentations as well as the well-established genetic modification tools available in a eukaryotic setting (Buntru et al., 2014, Buntru et al., 2015).

### Vectors

While the cellular extracts and lysates are used to create the CF expression platform, vectors are used to optimise it. Many features of vectors are designed and analysed for the optimisation of protein expression in the CF system, and are an important aspect of developing an efficient, high producing CF system that is qualified for being used in research. The features introduced in CF vectors have multiple origins and can come from species differing from those used to create the CF system (table 1.4). With the overall aim of improving protein expression, usually through limiting or abolishing bottlenecks in the transcription and translation, and the methods those features go by doing it varies. Commonly the aim of vector design is uniformity and compatibility across multiple CF expression systems. Ideally CF vectors have the potential for simple, quick cloning and can be used in pro- and eukaryotic systems to have high levels of protein production.

One common CF vector feature is the presence of IRES to aid in protein translation (Fitzgerald and Semler, 2009). Normally in typical protein translation the recruitment and binding of a 5’ cap on the mRNA is required for ribosomal recruitment and initiation of protein translation. In CF systems this demands the presence, in high volumes, of 5’ cap machinery and mRNA and therefore it acts as a bottleneck to the high-level production of proteins (Sonenberg and Hinnebusch, 2009, Brodel et al., 2013). Present within viral mRNA are nucleotide sequences that can bypass the cap-dependent protein translation protocol, known as IRES. These sequences, ranging in size from tens to hundreds of nucleotides, as the name suggests were discovered at the 5’ of poliovirus RNA (Pelletier and Sonenberg, 1988). Typically, the success of IRES usage in eukaryotes is species dependent, however there has been work carried out to develop a cross species functioning IRES. Multiple IRES sequences have been shown to work across multiple different eukaryotic species (Terenin et al., 2005) including the *Rhopalosiphum paid* viral IRES (Woolaway et al., 2001) and the Cricket paralysis viral IRES (Brodel et al., 2013) both shown to work in mammalian, plant and insect CF systems.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.4:** Vector elements utilised for improvement of protein transcription and translation, and increase in yields. | | | |
| **Feature** | **Origin** | **Usage** | **References** |
| **3’ UTR from viral RNA** | AMV | *In WGE*: role of 3’ UTR in translation initiation | (Ryabova et al., 1993) |
| **3’ UTR from viral RNA** | TMV | *In WGE*: Increased translation 3.5-fold through translational efficiency. TMV tail elongation by 160 nucleotides eliminated stimulatory effects. | (Zeyenko et al., 1994) |
| **5’ UTR from Ω gene** | TMV | *In WGE and ECE*: Increased translation of reporter proteins 2- to 10-fold. | (Gallie et al., 1987) |
| **5’ UTR from polyhedrin gene** | Baculovirus | *In ICE*: Improved protein translation. Vector containing 5’ UTR sequence showed 50-fold improved translational efficiency.  *In WGE and RRL*: Enhancement of translational activity. | (Ezure et al., 2006, Suzuki et al., 2006) |
| **5’ UTR from viral RNA** | EMCV | *In RRL*: Improved translation efficiency and initiation stringency. | (Craig et al., 1992) |
| **5’ UTR from IGR** | CrPV | *In ICE*: enhanced production of target proteins.  *In CHO extracts*: Cap independent translation, bypassing translation initiation bottleneck, enhancing translation levels. | (Brodel et al., 2013, Brodel et al., 2014) |
| **5’ poly(A) sequence** |  | *In RRL and WGE*: 25 adenylic residues before start codon acted as translational enhancer. Abolition of translational inhibition at high mRNA concentrations. | (Michel et al., 2000, Gudkov et al., 2005) |
| **5’ AT rich sequences** |  | *In ECE, WGE and RRL*: artificial AT rich sequences enhance protein expression across multiple CF systems. | (Mureev et al., 2009) |

With the increased usage and development of novel CF expression systems, the design and use of IRES-containing vectors has been an important aspect of it. Brodel et al. (2013) tested multiple IRES sequences from different viral origins across multiple eukaryotic CF systems with aim of increased expression that was tested with the glycoprotein erythropoietin and membrane proteins. Of the four IRES tested the CrPV IRES from the intergenic region showed the most increase in expression of luciferase protein along with EPO. The use of vectors containing IRES sequences has been used to develop CF expression systems including the CHO based system discussed earlier (Brodel et al., 2014). Other upstream sequences have been successfully used to improve CF systems (Carlson et al., 2012) including the use of omega (Ω) gene UTR from the tobacco mosaic virus (TMV) (Gallie, 2002), 3’ UTR from TMV RNA (Zeyenko et al., 1994) and 3’ UTR from alfalfa mosaic virus RNA 4 (Ryabova et al., 1993) in WGE CF systems, and the use of UTR of the polyhedrin gene from baculovirus (Suzuki et al., 2006). Outside of IRES, unique sequences have been shown to enhance protein translation, such as the poly(A) sequence tested in WGE (Gudkov et al., 2005). Species-independent translational sequences have been developed for use in CF systems (Mureev et al., 2009). These sequences, rich in poly-A or AT sequences, work in similar fashion to IRES by circumventing cap dependency for translation initiation in pro- and eukaryotic systems, and were successfully used for the development of a protozoan CF system from *Leishmania tarentolae* extracts. Using species-independent translational sequences containing vectors, high-throughput protein expression was achieved in prokaryotic and eukaryotic CF systems (Gagoski et al., 2015).

Other attempts on the use of vectors for optimising CF systems have been focused on the improvement of the effective plasmid concentration in CF systems. One system developed is the p-gel (protein-producing gel) system (Park et al., 2009). This system uses moulds to create hydrogel constructs that are subsequently covalently linked to plasmids creating a scaffolding system that protects the plasmid from degradation while also improving gene localization. Control of plasmid presence in CF system resulted in a 300-fold increase in protein expression levels in WGE and a 10-fold increase in produced protein activity. Similarly, a DNA bird nest hydrogel has been developed that is capable of hosting a much higher number of genes repeats (up to 32 000), while also offering the extra advantage of maintaining the link between genotype and phenotype (Kahn et al., 2016).

### Energy sources

Energy is an important driver of protein production and the more energy available, the higher the protein yields that can be achieved. Energy, in the form of ATP, is required by many processes in protein synthesis. Many methods have been developed that provide CF systems with energy, as well as offer the regeneration of it. The main source of ATP is the phosphorylation of ADP utilising high-energy phosphate bond containing molecules.

In *E. coli* a continuous fed-batch system was developed utilising phosphoenolpyruvate as the phosphate donor with pyruvate kinase catalysing its transformation into pyruvate (Kim and Swartz, 2000). With CF expression system, it was shown that PEP along with the amino acids arginine, cysteine and tryptophan were being depleted. Supplementation of the CF mixture with these components was found to increase protein synthesis with production time scale increased from 20 to 80 minutes. However, the increase in inorganic phosphate was found to inhibit protein production. To tackle this problem, Kim and Swartz (1999) developed a system using pyruvate oxidase from *Pediococcus* species to transform pyruvate into acetyl phosphate, this can be further transformed into acetate using acetyl kinase present in the *E. coli* extract which generated ATP molecules that can further fuel protein synthesis (Figure 1.6).

|  |
| --- |
|  |
| **Figure 1.6:** Methods for the generation of ATP driving cell-free transcription and translation |

Many other Sources for ATP used are glycolytic precursors of pyruvate, including glucose and glucose-6-phosphate (Calhoun and Swartz, 2005). PEP alternatives are cheaper to use and when used in combination with Bis-Tris buffer to stabilise pH levels, the level of chloramphenicol acetyltransferase protein synthesis is similar to that obtained with PEP as energy source. Fructose-1,6-bisphosphate was also shown to generate high levels of protein expression in *E. coli* while having the added benefit of not requiring additional cofactors such as NAD or Co-enzyme A (Kim et al., 2007). polymeric carbohydrates, with controlled glucose release, was used for longer protein expression periods was examined for decrease in organic acid accumulation and increase in pH values (Kim et al., 2011). The steady release of glucose from polymers prolonged CF expression for up to 12 hours and showed a much higher level of protein expression.

### Post-translational modifications

With protein expression systems, it is not enough to simply focus on going from DNA to protein with no concern towards the post-translational state of the protein. The usefulness of a protein in a process or in research is dependent on many factors that present after the translation stage of protein synthesis. Key aspects such as protein conformation and modification as well as purification and yield all have to be examined to enable CF expression systems to be utilised fully as legitimate sources of protein.

These PTMs are a vital finishing step in the synthesis of many proteins, especially those of more complex, eukaryotic origin. PTMs come in many different forms; from the attachment of chemical groups on the translated protein structure, such as glycosylation or phosphorylation, to the addition of hydrophobic lipid groups or the attachment of cofactors to the protein such as heme groups and flavin nucleotides. These modifications are important for the maintenance of the correct protein structure and are vital for the activation or inactivation of proteins. In cell-based protein expression systems the ability to carry out PTMs on proteins is dependent on the cell origin. Prokaryotic expression cells such as *E. coli*,or the eukaryotic yeast can carry out minimal PTMs, while eukaryotic expression cells such as insect and human cells are capable of producing more complex PTMs. This reflects the prevalence of PTMs in proteins of eukaryotic origin and its rarity in proteins of prokaryotic origin.

The presence of PTMs in many research and industrially valuable and important proteins shows the importance for creating CF protein expression systems capable of carrying out PTMs for expression platforms to be viable and to have large scale applications. Most of the research work focusing on potentials of CF systems capabilities of PTMs has been carried out in eukaryotic-based CF systems (Katzen and Kudlicki, 2006, Brodel et al., 2014), with some work focused on the potential of prokaryotic systems (Guarino and DeLisa, 2012).

Insect cells natively have the ability to carry out a limited range of PTMs. Many attempts in the literature have focused on the development of insect cellular lysates capable of protein production with PTM processing ability. A lot of effort has been on maintaining the glycosylation abilities of cellular lysates, with glycoproteins being widely present and play many vital roles in cellular function (Helenius and Aebi, 2004). Tarui et al. (2000) was able to successfully create an insect cell *S. frugiperda* Sf21 based CF protein expression system from a single step lysate extraction and synthesise the glycoprotein gp120 in the correct conformational structure and N-linked glycosylation modifications capable of binding to CD4 proteins. The lysate extraction method used was a Mini-Bomb cell disruption chamber under nitrogen gas conditions. Another method of insect cell lysate preparation used a French pressure cell press to carry out the cellular disruption (Katzen and Kudlicki, 2006). The glycosylation capability of this Sf21 insect CF system lysate was shown to be better in comparison to lysates prepared using the Mini-Bomb cell disruption chamber. In addition, the lysate was shown to be efficient in the cleaving of the signal sequence of the β-lactamase protein. With ubiquitination in insect cell lysates, Suzuki et al. (2010) were able through mass spectrometry confirm the capability of insect cell lysates to carry out poly-ubiquitination of the p53 protein. The CF reaction mixture was prepared with commercial cell extracts, with added ubiquitin, ubiquitin aldehyde (a de-ubiquitination inhibitor) and Mdm2 (an ubiquitin ligase).

The addition of foreign elements into CF expression mixtures is commonly utilized for PTM processing. Early attempts at improving glycosylation in CF expression systems involved around the addition of microsomes; membrane vesicles originating from the endoplasmic reticulum and are the glycosylating tools of cells. Usually of canine origin, microsomes have been added to cell lysates of wheat germ (Lingappa et al., 1978) or rabbit reticulocytes (Walter and Blobel, 1983). However, these attempts were limited in success and proved costly. For glycosylation of proteins expressed in human hybridoma cell lysates, the addition of transcription factors GADD34 and K3L improved glycosylation of gp120 protein (Mikami et al., 2006a). In that study, the hybridoma cells were disrupted using the Mini-bomb cell disruption chamber. The addition of GADD34 and K3L lowers the phosphorylation of eIFα, a eukaryotic translation initiation factor. This combats the elevated levels of phosphorylation observed by the addition of the essential ATP and creatine phosphate supplements of CF reaction mixtures. PTMs other than glycosylation are also very important and have been looked at in detail, including acetylation (Rubenstein et al., 1981) and phosphorylation (Joshi et al., 1995).

While traditionally most of the CF PTMs work has been focused in eukaryotic systems, there has recently been a push towards demonstrating PTMs capabilities in prokaryotic systems. Guarino and DeLisa (2012) have utilized the N-linked glycosylation pathways of the *Campylobacter jejuni* bacteria. The pathway utilizes the *pgl* gene cluster with the resultant glycosylation similar to that of eukaryotes (Glover et al., 2005b). The gene cluster has previously been transformed into *E. coli* cells and successfully performed N-linked glycosylation (Wacker et al., 2002). In the CF study, the AcrA protein with two glycosylation sites was utilized as the model protein. It was found that the AcrA was produced predominantly in its twice glycosylated variant, with small amounts of AcrA present in it singularly glycosylated variant when the protein was expressed in two different *in vitro* systems, *E. coli* CF system and PURE system. In both systems, the glycosylation took place in the presence of *Campylobacter jejuni* oligosaccharyl transferase PglB, lipid-linked oligosaccharide and the acceptor protein. Looking beyond the functions of PglB in CF systems, Glover et al. (2005a) have looked into the reconstitution of other oligosaccharyl transferases, such as PglA, J, H and lI, opening a possible use in future CF systems for more complex PTMs.

The disulphide bond is a covalent bond between two sulphur atoms of cysteine amino acids. It is one of the most important protein bonds, playing a role in the structure and stability proteins. With the applications of CF expression systems being explored, for it to be successfully used in structural protein analysis, correct protein bond formation and folding is required. The challenge with disulphide bonds is the creation of an environment conducive to the formation of the bond. The desired environment for disulphide bond formation is an oxidising one. The environments of bacterial and eukaryotic cells are different and therefore the localisations and machinery for disulphide bond formation is different in both.

In Gram-negative bacterial cells, such as *E. coli*, the cytosol is a reducing environment unsuitable for disulphide bond formation; instead the bonds are catalysed on the cellular membrane in the periplasm (Berkmen, 2012). When creating bacterial CF systems, the lysate environment has to be altered to favour the formation of disulphide bonds. The main focus was on the control of the redox potential of the lysate. Yin and Swartz (2004) carried out the expression of plasminogen activator protein with multiple disulphide bonds utilising an *E. coli* CF system. For the improvement in the disulphide bond formation, the lysate was supplemented with iodoacetamide in addition to a mix of oxidised and reduced glutathione to create an optimised redox potential. To further increase expression levels DsbC (disulphide bond isomerase) and Skp (*E. coli* periplasmic chaperone) were added to the reaction lysate. Further work by Knapp et al. (2007) aimed to tackle the reduction of oxidised glutathione caused by the glutathione reductase and thioredoxin reductase using a lower iodoacetamide concentration, allowing the use of glucose as a carbon source. Another study utilised L-buthionine sulfoximine, an inhibitor of glutamate-cysteine ligase to inhibit the formation of reduced glutathione and preserve the ratio of oxidised to reduced glutathione added to *E. coli* CF lysates (Oh et al., 2007).

Eukaryotic Sf21 insect cells were used to express the multiple disulphide bond containing alkaline phosphatase and human lysozyme proteins. In similar protocol to the bacterial CF system preparation, a mix of oxidised and reduced glutathione was added to the CF lysates, with an additional protein disulphide isomerase enzyme (Ezure et al., 2007). The addition of protein disulphide isomerase has been used elsewhere to improve disulphide bond formation in both in *E. coli* (Ryabova et al., 1997) and WGE (Kawasaki et al., 2003).

### Commercial CF systems

Many companies have noticed the benefits CF expression systems offer in research and have taken advantage of it by offering ready-prepared kits for CF protein expression. The preparation of CF extracts can be laborious, complicated and time consuming. The commercial CF expression systems are pre-prepared and shipped to the lab ready for use, often only requiring mixing of components (*e.g.* addition of amino acids, energy sources and labelling molecules to cell extracts), the addition of template DNA or mRNA and the incubation for a set amount of time. There is a wide range of commercially available CF protein expression systems now (Table 1.5).

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.5:** Commercially available CF protein expression systems and their key features. | | | |
| **Origin** | **CF system** | **Manufacturer** | **Description** |
| ***E. coli*** | PURExpress In Vitro Protein Synthesis Kit | NEB | Reconstituted from the purified components necessary for *E. coli* translation. |
| S30 T7 High-Yield Protein Expression System | Promega | Quick, high-yield expression, up to 500 µg/ml.  T7 RNA polymerase containing extract, deficient in OmpT endoproteinase and lon protease activity. |
| **Insect** | TNT T7 Insect Cell Extract Protein Expression System | Promega | Sf21 *Spodoptera frugiperda*  Protein is expressed downstream of the T7 RNA polymerase promoter. Companion vectors were designed to achieve optimal yield with this system. |
| **Wheat germ** | TnT SP6 High-Yield Wheat Germ Protein Expression System | Promega | Yields of 10-100 µg/ml.  Expresses genes downstream from an SP6 RNA polymerase promoter. |
| **Rabbit reticulocyte** | TNT Quick Coupled Transcription/  Translation System | Promega | Genes cloned downstream from either the T7, T3 or SP6 RNA polymerase promoters. |
| **Human** | 1-Step Human coupled IVT Kit | Life Technologies | Yields of up to 100 µg/ml.  Optimized pT7CFE1 Expression Vector.  Human translational machinery and PTMs. |
| Human Cell-Free Protein Expression System | Takara Bio | Dialysis system for ATP and amino acid replenishment.  T7 vectors with IRES.  Human translational machinery and PTMs. |
| **Chinese Hamster Ovary** | 1-Step CHO High Yield IVT Kit | Life Technologies | High yield up to 750 µg/ml.  Continuous-exchange CF system. |

Each commercial product is designed for targeted functions where it provides the optimal results. The CF systems have diverse origins, covering both prokaryotic and eukaryotic cell lysates origins. These systems are advertised for their simplicity and efficiency, with protein expression possible within the time frame of a few of hours rather than days and weeks that is required with cell-based protein expression.

### Applications

CF systems have a wide range of applications in research (figure 1.5). CF systems have been modified to include and utilise unnatural amino acids aiding with the development of proteins with novel functions and structures. The expression of toxic proteins in CF systems has allowed the exploration of many disease-causing proteins and the development of treatments. Other difficult to express proteins such as membrane proteins that are hydrophobic and have a tendency towards aggregation and inclusion body formation have been successfully expressed, including on the surfaces of lipid and polymer vesicles. The vast field of synthetic biology has also utilised CF systems fully including for the development of minimal cells, *in vitro* compartmentalization and the development of novel production pathways.

#### Unnatural amino acids

There are many additional amino acids outside of the twenty canonical that can be utilised as building blocks for proteins, however since organisms *in vivo* do not possess the machinery capable of incorporating these non-canonical amino acids into a protein sequence, these uAAs are difficult to incorporate into protein sequences. There are many advantages to the utilisation of uAAs (*e.g.* structural studies, protein engineering and the creation of novel function proteins) that drives the development of uAAs incorporation techniques (Zhang et al., 2013). In *In vivo* expression, there have been difficulties faced in incorporating uAAs into expressed proteins. The common strategy utilised is the orthogonal synthetase and tRNA pairing, previously carried out in bacterial and eukaryotic cells (Neumann, 2012). These rely on the availability of the uAA in the system, an unused codon, a tRNA that can recognise that codon and a tRNA synthetase that can recognise both the tRNA and the uAA. Typically, the unused codon is a reassigned stop codon or by creating a four-base codon. This has been carried out successfully but is still limited with only one uAA incorporated into a protein at a time. *In vitro* systems, with its ease of control and modification, it offers new potentials into the expansion of the genetic code. CF protein expression systems have been utilised with a modified genetic code where codons code for tRNAs carrying uAAs and ribosomes that can attach them to the AA sequence.

Early incorporation of uAAs in CF systems was carried out using *E. coli* S30 extracts (Noren et al., 1989, Kanda et al., 2000). Methods of introducing the uAAs in CF systems include the utilisation of chemically acylated suppressor tRNA binding to the stop codon substituted into the position of interest (Noren et al., 1989), utilisation of a four/five nucleotide anti-codon recognition tRNA (Hohsaka et al., 2001a, Hohsaka et al., 2001b) and the utilisation of anti-sense DNA treated *E. coli*  extract for the inactivation of tRNAs and their substitution into uAAs (Kanda et al., 2000). High-yield protocols have been developed specifically for *E. coli* site-specific uAAs incorporation, using aminoacyl-tRNA synthetase/suppressor tRNA pairs derived from *Methanocaldococcus jannaschii* (Ozawa et al., 2012).

#### Toxic proteins

There are a wide range of toxic proteins that causes the death of cells they are expressed and present in, so called cytotoxic proteins. They are present in a wide range of species and fill a wide range of roles. Typical cytotoxic proteins include those present in disease states or originating from disease causing organisms. Their study is vital for understanding of cellular mechanisms yet their expression is a challenge hindering progress. CF systems offer an alternative for high level protein yields without any hindrance. Examples of cytotoxic proteins synthesised in *E. coli* CF systems include the membrane proteins subunit *b* of F1F0 ATP synthetase after a simple four hour expression (Lian et al., 2009) and the adenine nucleotide translocator for the development of a drug discovery model (Yang et al., 2012). GFP has been fused to the human cytotoxic peptides β-defensin-3 and 4 in an *E. coli* CF system for quantitative monitoring of protein synthesis levels (Chen et al., 2007). A group of subunits of the bacterial cytolethal distending toxin cytotoxic proteins were expressed using an *E. coli* CF system and through the utilisation of mutagenesis, the functions of the different subunits were elucidated (Avenaud et al., 2004).

CF systems have been examined to see the effect of protein synthesis inhibiting toxins on protein expression. *E. coli* toxinYoeB is a translation initiation inhibitor that was shown to inhibit expression of proteins in prokaryotic CF system, with the YefM antitoxin restoring protein expression (Zhang and Inouye, 2009). No YoeB inhibitory effect was monitored in eukaryotic CF expression systems. The prokaryotic CF system was used to show which cellular components were important for toxin function, YoeB did not exhibit its endoribonuclease activity until it was present with 70 S ribosomes, suggesting that it is a ribosome-associating factor.

#### Membrane proteins

Membrane proteins are a vital group of proteins and of special interest to the pharmaceutical industry. Expression of membrane proteins due to their insolubility can be troublesome within cells. Cellular machinery is often ill equipped to process over-expression of hydrophobic proteins. Such proteins tend to be misfolded, form insoluble inclusion body aggregates of non-functional proteins and lead to cell death. CF expression has emerged as an alternative with great potential for the functional expression of membrane proteins (Schwarz et al., 2008, Rajesh et al., 2011). When utilising CF systems, the proteins expressed are exposed and open for interaction, it is possible to add components to the system that interact with the proteins and allow for increased expression, improved solubility and functionality. Wide ranges of detergents, membrane mimetics and lipid particles have all been used in conjunction with CF systems with that aim. This will be useful in the study in microsomal monooxygenases as it allows for the direct supplementation of the required cofactors, including the heme precursors of δ-Aminolevulinic acid or hemin for microsomal monooxygenases and C*b5* as both contain and require heme for functionality, as well as the FAD and FMN needed for CPR functionality.

Structures such as liposomes, polymersomes and nanodiscs can act as biomimetic membranes. The aim of these components is to provide membrane proteins a hydrophobic stabilising environment similar to their native environment that is necessary for protein folding and function. Exposing membrane proteins as they are being translated to these environments prevents any undesired aggregation or misfolding. Proteoliposomes, liposomes with inserted proteins on the surface, have been successfully used to study protein function (Sachse et al., 2014). Similarly, polymersomes are used in a similar fashion to liposomes as a stage for membrane protein insertion. There has been an increase in the usage of polymersomes over liposomes in research due to their ease of modification, improved robustness and stability. Multiple different membrane proteins kinds have been successfully inserted into polymersome surface, creating so called proteopolymersomes. Outside of liposome and polymersome, there have been studies on the effect of utilisation of hydrophobic supplements (*e.g.* detergents, amphipols, peptide surfactants and bicelles) for CF systems (Rajesh et al., 2011, Hein et al., 2014). There have been multiple studies to see the effects of detergent use on CF expression system and membrane protein expression. Typically, most detergents tested have shown very limited interference with transcription and translation machinery. One detergent, Brij-35, was tested against 13 GPCRs from different classes showing correct folding and ligand binding when proteins were expressed in an *E. coli* CF system (Corin et al., 2011). Native membrane vesicles have been favoured for MP production in CF systems. *E. coli* membrane vesicles have been added to *E. coli* based CF system and has shown successful production and folding of two membrane proteins (Wuu and Swartz, 2008). In eukaryotic systems, microsomes supplemented Insect Sf21 lysates have been used to produce correctly folded and functional membrane channel proteins (Quast et al., 2015). Nanodiscs, created by Bayburt et al. (1998), are a phospholipid bilayer held in a disc arrangement by membrane scaffold proteins. These mini membrane structures have been used in conjunction with CF system for membrane protein hosting, including cytochrome P450 enzymes and GPCRs (Yang et al., 2011) and integral membrane proteins (Lyukmanova et al., 2012).

#### Synthetic biology

Synthetic biology is an ever-increasing research field with more and more work being done taking advantage of the principles synthetic biology to tackle some of the biggest biological problems. CF systems offer unique advantages as a platform for synthetic biology, that are simply not available nor feasible with the commonly utilised cell-based platforms, helping expand the toolbox available for synthetic biology. CF systems have been combined with synthetic biology for the development of protein and metabolite production, regulatory circuits, compartmentalization and minimal cells (Hodgman and Jewett, 2012).

With CF systems, each component can be carefully purified, selected and combined to carry out a set function. It is possible to create manufacturing platforms for the production of high value metabolites, *e.g.* biofuels and biopharmaceuticals. Although as of yet, CF systems cannot compete at the level of the traditional manufacturing organisms, it is improving. *E. coli* CF systems have been used for the production of bioactive therapeutics (Kim and Swartz, 2004), and virus-like particles (Bundy et al., 2008) and scFv fusion proteins (Kanter et al., 2007) for vaccine development. Similarly VLPs have been synthesised in yeast CF system (Wang et al., 2008). virus-like particles in particular show great potential with CF systems due to the limited control available in *in vivo* systems, VLP toxicity and effects on cell growth (Rodríguez-Limas et al., 2013, Zeltins, 2013), and this potential is highlighted by the production of virus-like particles containing toxic proteins (Smith et al., 2012) and uAAs (Patel and Swartz, 2011). Difficult to express *in vivo* proteins such as [FeFe] hydrogenases, due to their complexity and oxygen sensitivity, have been produced using *E. coli* CF systems supplemented with three proteins from *Shewanella oneidensis* (Boyer et al., 2008). To the benefit pharmaceutical production, there has been research into the developing high throughput technology with CF systems for screening, expression and engineering of proteins (Casteleijn et al., 2013, Murray and Baliga, 2013).

Fine control of expression, beyond common bacteriophage regulatory elements, has been shown to be possible in CF systems with the development of multiple types of genetic and transcriptional circuitry, including multi-step expression cascades, transcriptional oscillators and ribosomal switches (Hodgman and Jewett, 2012). Multi-step expression cascades work on the principle that a promoter is activated and in turn expressing factors that regulate further expression. Noireaux et al. (2003) developed a two and three step cascade circuit with both transcriptional activation and suppression. The RNA polymerases (T7 and SP6) were activated in series to express the luciferase protein, and control of expression was shown, however there was time delay and decrease in output attributed to the bottleneck at the transcriptional machinery (Noireaux et al., 2003). Since then more complex *in vitro* circuits have been developed in *E. coli* CF system containing multiple stage cascades, gates and negative feedback loops (Shin and Noireaux, 2012) and new prototyping of linear circuits using *E. coli* extracts in the newly developed transcription-translation system (TX-TL) (Sun et al., 2014). Development of biological *in vitro* transcriptional oscillators has been carried out with negative feedback, positive feedback and three switch ring oscillators using RNA molecule signals (Kim and Winfree, 2011). Regulatory riboswitch elements offer up a valuable new tool in the push for control of protein expression (Muranaka and Yokobayashi, 2010). CF circuitry has been thoroughly reviewed (Hockenberry and Jewett, 2012, Hodgman and Jewett, 2012, Weitz and Simmel, 2012) and has led to increasing our understanding of CF system expression in terms of characterising the kinetics and flux through the system with the overall aim of expression optimisation.

*In vitro* compartmentalization, developed by Tawfik and Griffiths (1998), links a set genotype to its phenotype in a defined environment. This technique has proven incredible in the evolution of a wide range of enzymes, including phosphoesterases (Griffiths and Tawfik, 2003) and [FeFe] hydrogenases (Stapleton and Swartz, 2010). Through this directed evolution activity, specificity, tolerance and stability of enzymes have been modified and improved. Within IVC CF expression plays a key role in the transcription and translation of the genetic information and presenting an assayable phenotype. It is desirable to use expression systems that are fully capable of producing the active enzyme, *i.e.* express eukaryotic enzymes using eukaryotic CF extracts. Currently CF systems from bacterial (Griffiths and Tawfik, 2003) and eukaryotic (Doi et al., 2004, Ghadessy and Holliger, 2004) origins have been used in combination with *In vitro* compartmentalization.

Minimal cells, like *In vitro* compartmentalization, have been developed through the compartmentalization of *in vitro* transcription and translation mixtures for the purpose of protein expression. The aim of these so called minimal cells is to create cell mimicking compartment structures self-sufficient in protein expression and replication. The components within these cells should only be the minimal requirements for it to carry out its functions; transcription, translation and replication, hence their minimal cell nature. The bottom-up *in vitro* approach for minimal cells uses key components for protein expression along with DNA and RNA compartmentalized in membranes (Szostak et al., 2001). Being a functional cell and capable of programmed replication has presented its own advantages including the creation of a synthetic biological unit that has applications in protein and metabolite production, as well as understanding cellular networks. Challenges of complexity (Jewett and Forster, 2010, Stano and Luisi, 2013) and crowding (Tan et al., 2013, Sokolova et al., 2013) are faced and tackled in the development of minimal cells. Semi-synthetic minimal cells, with active protein expression and functional membrane components have been developed using liposomes (Nourian et al., 2012) and polymersomes (Martino et al., 2012).

One recent advancement is iSAT (integrated synthesis, assembly and integration), which aims to create a minimal cell like system where rRNA is synthesized utilizing native ribosomal proteins from cellular extracts, coupled with their assembly and followed by the synthesis of functional proteins (Jewett et al., 2013). In this study, ribosome free S150 *E. coli* extracts were used to construct clindamycin resistant mutant 23 S rRNA. Further study of the iSAT system, aimed at improving its efficiency through the control of its transcriptional activity has been carried out. Fritz and Jewett (2014) utilized GFP as a model protein improvement of iSAT, achieving a 45-fold increase in protein production. This improvement in protein production rates was achieved through 3’ modifications in the gene sequences of the rRNA, alteration of the concentration of plasmids and polymerases, and the use of an rRNA operon under the control of a T7 promoter that is able to balance rRNA synthesis and processing.

Similar to IVC, mRNA and ribosome display are used to link a desired protein phenotype with the responsible mRNA for the purposes of directed evolution of proteins, however importantly it is distinguished from IVC due to the absence of compartmentalization (Murray and Baliga, 2013). Use of directed evolution is a key tool in the altering and optimisation of proteins for tailored applications with important potential for pharmaceutical production of proteins (Tee and Wong, 2013). In ribosomal display, transcribed RNA from a DNA library is stalled and tethered at the ribosome (due to the absence of stop codon activating release factors, low temperatures and high concentration of Mg2+ ions) while the protein sequence protrudes sufficiently from the ribosome for folding into its active conformation (Hanes and Pluckthun, 1997). Proteins and peptides can be screened and selected at high rates (with *in vitro* ribosomal display the limiting factors are the number of ribosomes, in comparison to *in vivo* displays where transformation efficiency is a hindrance). Similarly, in mRNA display, mRNA sequences are covalently ligated to adaptor molecules that with translation link the mRNA to the peptide sequence. These desired sequences can be purified, reverse transcribed and amplified or recycled into multiple rounds of display evolution (He and Taussig, 2002). This technology has proved highly valuable for the development of important therapeutics and diagnostics through the screening and study of antibodies (Murray and Baliga, 2013).

### Future outlooks

CF *in vitro* protein expression systems offer great benefits in research. A wide range of applications in research, often uniquely only offered through CF systems has kept these systems at the forefront of research. The analysis and optimisation of every aspect and component of CF systems has been a vital process in the improving the output offered by CF systems, with optimisation of lysate harvesting, vector design and energy provision increasing protein yields significantly. There is still continuous research being carried out looking to improve the capabilities of CF even further, as well as developing tools and techniques to aid in the future development of novel CF systems and optimisation of current systems. Technologies like SITS vector elements and their broad species functionality offer a great resource in the optimisation of CF systems, and the similar exploration of other technologies, through scouting for features in nature, with different optimisation method are sought after to actively bring CF expression in level with *in vivo* expression. The development of novel CF systems, such as the CHO and the TBY2 extracts are important for the diversification of CF systems on offer beyond the current major systems which will cater to specialised research requirements and widen the scope of CF applications. Additionally, the advancements in CF systems preparation protocols cannot be understated, breakthroughs such as those seen in the time and methodology of CF extract preparation, going from a costly multiple day protocol to just a couple of hours and utilising common lab equipment greatly increases their affordability and brings their use into more and more labs. Creating truly streamlined CF expression platforms can and will be capable of competing and even outperforming the traditional cell-based protein expression.

One key functionality of CF systems that will be the main drive behind most of future work utilising it is that it is an open system that allows for direct interaction with the system and expressed proteins, offer greater degree of control and tunability of the expression environment. This potential of CF expression systems is of importance in the study of microsomal monooxygenases, and cannot be underestimated. Since the development of artificial biomimetic cell membranes, they have been used for the study of wide range of membrane-bound proteins, including a wide range of cytochrome P450 enzymes. While most research utilising these biomimetic membranes have involved cell expressed proteins; more and more recent work has been utilising CF protein expression systems.

## Biomimetic cell membranes

### Introduction

The value of research into microsomal monooxygenases has been made clear, they are of immense importance to the understanding of the metabolism of most drugs available on the market, prediction and understanding of drug side-effects and designing of future drug targets including individualised and patient driven design. CF expression systems have also been shown to be a highly useful tool for the expression of the microsomal monooxygenases and indeed they open new avenues in the research of these enzymes that was previously hampered by the standard cell-based expression systems. These CF systems have greatly simplified the potential usage of biomimetic membranes in membrane protein usage, and although limited research has been done utilising these biomimetics in understanding microsomal monooxygenases, there is a lot of potential for further work.

Already various types of native and biomimetic membranes have been utilised for the expression of membrane binding proteins to provide the holding and stabilising environment that allows for protein folding and its presentation in the most ideal conformation for function. They can tackle the formation of aggregates that is often observed with the over-expression of these membrane proteins in the absence of suitable membranes, leading to increased stress on the expression systems and increased toxicity. The aggregates themselves are usually misfolded, have little to no function and are difficult to purify and reconstitute. Alterations in the expression strategies of these proteins can however bypass these obstacles, namely the utilisation of native membranes already present within the expression system or supplementation of artificial membranes that can carry out such function. Further modifications of expression systems can be used to optimise protein insertion and stabilisation into the membrane including the use of chaperones, signalling sequences and detergents.

The most commonly utilised membranes have been the native membranes present within the expression systems themselves, including the cell membranes and membranes within the cell such as the endoplasmic reticulum and mitochondria. Protein expression looking to utilise such membranes usually require complex peptide targeting mechanisms like signal peptides and chaperones for post-translational delivery to the desired membrane localisation. However, biomimetic membranes can offer a simpler alternative for the stabilisation and solubilisation of these challenging proteins.

### Biomimetic membranes

Biomimetic membranes have been developed to help research and better understand cellular membranes and their surrounding environment, from the proteins embedded within them to their interactions with molecules and environment. Very broad and varying types of biomimetic membranes have been developed based on lipid and polymer building blocks.

The classic example of lipid-based membranes, the vesicular liposome, utilises lipids composed of hydrophilic “heads” and hydrophobic “tails” that can rearrange into the bilayer arrangement seen in native membranes. Various methods can be utilised to force the lipids into a vesicular formation, including extrusion, electroformation and sonication. Fine tuning of liposome can be carried out controlling all characteristics such as lipids used, size, formation solvent and membrane components. Alternative lipid based biomimetic membranes have been also been developed including nanodiscs and planar lipid membranes. Nanodiscs are composed of bilayer lipid structures held together by membrane scaffolding proteins (Bayburt et al., 1998). Liposomes and nanodiscs have been used as biomimetic membranes, providing membrane proteins the hydrophobic stabilising environment similar to their native environment that is necessary for protein folding and function.

Proteoliposomes have been successfully used to study protein function (Sachse et al., 2014), the predominant method for the production of proteoliposomes is through protein reconstitution methods utilising cell-based expression systems (Rigaud and Levy, 2003). It has successfully been utilised for all membrane protein classes including receptors, enzymes and channel proteins. Using CF systems, exposing membrane proteins as they are being translated to these lipid environments can prevent any undesired aggregation or misfolding. Proteoliposomes containing the CF expressed integral membrane protein connexion-43 were created and shown to be functional through the delivery of small hydrophilic peptides across the liposome membrane using the rabbit reticulocyte CF system (Kaneda et al., 2009), while connexion-43 expressed in an *E. coli* PURE system showed protein integration in an oligomeric fashion (Moritani et al., 2010). Characterisation studies of membrane channels were shown to be possible using *E. coli* CF systems, the *E. coli* water transporter aquaporin Z was successfully expressed using S30 *E. coli* lysate system and functional activity was observed (Hovijitra et al., 2009). There have been other classes of membrane proteins successfully inserted into liposome membranes, including cytochrome enzymes expressed in a wheat germ system (Nomura et al., 2008), *E. coli* MscL channel without use of insertase proteins in an *E. coli* CF system (Berrier et al., 2011) and *E. coli* expressed bacteriorhodopsin proton pumps (Kalmbach et al., 2007). The presence of liposome has been shown to affect CF protein expression levels (Bui et al., 2008).

### Polymersomes

Amphiphilic polymers can self-assemble in aqueous solutions into ordered structures including biomimetic polymersome vesicles (Discher et al., 1999). Di- (AB) and tri- (ABA) block copolymers have been used as the building blocks of vesicles with monolayer membranes (table 1.6, figure 1.7). Similarly, to other biomimetic membranes, they have been used as a platform for membrane protein insertion (Nallani et al., 2011, May et al., 2013), as well as for encapsulation of drugs and other compounds of interest (Arifin and Palmer, 2005, Rameez et al., 2008, Messager et al., 2014).

|  |  |  |
| --- | --- | --- |
| **Table 1.6**: Examples of polymers utilised in the synthesis of polymersomes | | |
| **Arrangement** | **Polymer-*block­*-Polymer** | **Ref.** | |
| **AB**  **(Diblock)** | Poly(butadiene)-*b*-Poly(ethylene oxide) *(PBD-PEO)* | (Arifin and Palmer, 2005, Li et al., 2007, May et al., 2013) | |
| Poly(lactic acid)-*b*-Poly(ethylene oxide) *(PLA-PEO)* | (Domes et al., 2010, Leclercq and Vert, 2017, Xiao et al., 2010) | |
| Poly(dimethylsiloxane)-*b*-Poly(2-methyloxazoline)  *(PDMS-PMOXA)* | (Egli et al., 2011a) | |
| **ABA**  **(Triblock)** | Poly(2-methyloxazoline)-*b*-Poly(dimethylsiloxane)-*b*-Poly(2-methyloxazoline)  *(PMOXA-PDMS-PMOXA)* | (Nardin et al., 2000, Fu et al., 2011) | |
| Poly(ethylene oxide)-*b*-poly(oxypropylene)-*b*-Poly(ethylene oxide) *(PEO-PPO-PEO)* | (Rodríguez-García et al., 2011) | |
| Poly(ethylene oxide)-*b*- poly(dimethylsiloxane)-*b*-Poly(ethylene oxide) *(PEO-PDMS-PEO)* | (Choi and Montemagno, 2005, Choi et al., 2006) | |
| **ABC**  **(Triblock)** | Poly(ethylene oxide)-*b*-Poly(dimethylsiloxane)-*b*-Poly(2-methyloxazoline)  *(PEO-PDMS-PMOXA)* | (Stoenescu et al., 2004) | |

|  |
| --- |
|  |
| **Figure 1.7:** Arrangement of the hydrophilic and hydrophobic regions on polymers, and their arrangement on vesicle membranes. The diblock (AB) membrane arrangement mimics that of the lipid bilayer with hydrophobic regions aligning in the middle of the membrane and the hydrophilic regions are arranged outwards. In the triblock (ABA/ABC) membrane arrangement the hydrophobic region is buried within the polymer membrane and surrounded by hydrophilic polymers. |

There has been an increase in the usage of polymersomes over liposomes in research due to their ease of modification, improved robustness and stability, decreased levels of permeability and longer shelf-life (Discher et al., 1999, Discher et al., 2007, Discher and Ahmed, 2006). The chemical diversity of available polymers offers unlimited variability and tunability of polymersomes, from membrane thickness to hydrophobicity, permeability and rigidity can all be altered and tuned for purpose. The lateral mobility of polymers and membrane viscosity has been shown to decrease in polymersomes composed of higher molecular weight polymers (Discher and Ahmed, 2006). The decrease in polymersome fluidity is caused by the entanglement of the longer length polymers. Similarly, a decrease of permeability through the membrane is observed with higher molecular weight polymers. The lack of polymersome permeability is not a rigid characteristic and the tuning of permeability for specific functions has been carried out using hybrid polymer-lipid vesicles, showing increased permeability with increased lipid presence (Lim et al., 2013). The creation of hybrid diblock and more permeable triblock copolymer polymersomes has also shown to increase membrane permeability (Rodríguez-García et al., 2011, Yan et al., 2015), and was similarly attempted with two different diblock copolymers (Gaspard et al., 2016). Functionalization of polymersomes for desired applications through conjugation to preformed polymersomes, or through utilisation of functionalised polymers for polymersome assembly is possible for a variable range of ligands (Egli et al., 2011b), including with antibodies, adhesin compounds and fluorescent molecules (*e.g.* rhodamine B) conjugated to lipid moieties that can stably insert into polymersome membranes .

#### Stability and integrity

Polymersomes are desired for their superior levels of structural stability and integrity compared to lipid based membranes (Discher et al., 1999, Lee et al., 2001, Rodríguez-García et al., 2011). The vast chemistry of polymers available allows for the design of polymersomes with desired levels of stability and rigidity. The simplest method of changing the rigidity of the membrane is by altering molecular weight of the polymer, through the alteration of the polymer length. Standard lipids are short, have low molecular weights, and membrane hydrophobic core thickness of around 3-5 nm, while on the other hand polymersomes, with variable polymer lengths have shown it is possible to much broader core membrane thickness of 8-21 nm (Discher et al., 2007). The thick and tightly wound up polymer arrangement can form a rigid and impermeable barrier that maintains the vesicle stability over longer periods, and through creating vesicles with polymers having longer hydrophobic regions, the formed polymersomes showed increased tension and strain limits (Bermudez et al., 2001). Modification of polymersomes for increased permeability has also been shown to be possible. Through the usage of triblock copolymers composed of the moderately hydrophobic oxypropelyne flanked by the hydrophilic polyethylene oxide, this overall relatively hydrophilic polymer showed increased levels of permeability in comparison to standard PBD-PEO polymersomes (Rodríguez-García et al., 2011).

On the other end of the spectrum, it’s possible also to create biodegradable polymersomes. These polymersomes have great potential as carrier vesicles that with targeted and controlled release. PLA polymers have shown susceptibility towards hydrolytic degradation (Batycky et al., 1997), and although independently their use has been hampered by its weak hydrophilic characteristics, prolonged degradation time and low uptake of polar compounds, through combination with PEO polymers, it is possible to alter and improve its biodegrading and biocompatibility characteristics (Xiao et al., 2010). To expand the array of degradable polymers available, other biocompatible polymers have been synthesised including PCL based diblock copolymers (Katz et al., 2010).

### Proteopolymersomes

The production of protein carrying polymersomes, so called proteopolymersomes, is still in the early development stage. Various techniques have been used for the incorporation of proteins on the membrane surface with varying degrees of success, and varying success has been achieved with the isolation of these proteopolymersomes for further downstream applications.

#### Protein incorporation

Strategies for the incorporation of proteins into polymersomes fall into two main categories; reconstitution through detergent removal, and co-translational insertion, and both have been regularly utilised (table 1.7). Reconstitution protocols typically utilise cell-based protein expression systems, *e.g. E. coli*, yeast, insect cells or endogenous cells for their high protein yields. These cells are lysed and their proteins are purified in detergent containing buffer. This buffer allows the purification of membrane proteins in the form of stable protein-detergent complexes. Proteins present within native membranes are targeted for destabilization gradually by detergents, firstly detergent-lipid complexes are formed and eventually with sufficient detergent, the lipids are destabilised to form the detergent-protein complex. Gradual removal of detergents in the presence of polymersomes drives the gradual transfer of protein into the polymersome membrane. Issues with the solubilisation of membrane proteins using detergents can occur, and is not always guaranteed (Seddon et al., 2004). There are three classes of detergents; Ionic (cationic or anionic), non-ionic or zwitterionic, based on the characteristics of the polar “head” group of the detergent. Each class generally has its own advantages and disadvantages which affect its utilisation. Ionic detergents (*e.g.* Sodium dodecyl sulphate) can have higher degrees of solubilisation but also a higher risk of protein denaturing, in comparison non-ionic detergents (*e.g.* dodecyl maltoside) are more mild but lower solubilisation capacities. Zwitterionic detergents (*e.g.* FOS-choline 14) combine characteristics from both ionic and non-ionic detergents.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.7:** Examples of proteopolymersome production | | | |
| **Insertion method** | **Protein** | **Polymersome** | **Ref.** |
| **Reconstitution** | Aquaporin Z | Triblock  (PMOXA-PDMS-PMOXA) | (Kumar et al., 2007) |
| ATP synthase | Triblock  (PEO-PDMS-PEO) | (Choi and Montemagno, 2005) |
| MloK potassium channel | Diblock (PDMS-PMOXA) | (Kowal et al., 2014) |
| **Co-translational** | Dopamine D2 receptor | Di- (PBD-PEO) and triblock (PMOXA-PDMS-PMOXA) | (May et al., 2013) |
| Claudin 2 | Diblock (PBD-PEO) | (Nallani et al., 2011) |
| Light harvesting complex II | Diblock (PBD-PEO) | (Zapf et al., 2015) |
| Chemokine receptor 4 | Diblock (PBD-PEO) | (de Hoog et al., 2014) |

Generally, the best detergent is dependent on the protein of interest; for example short-chain alkyl maltosides have been most commonly used for the purification of membrane proteins (Nollert, 2005) while longer chain maltosides have been used in the purification of oligomeric GPCRs (Jastrzebska et al., 2006). Detergent screening is usually carried out for individual proteins, and solubilisation can be monitored by assaying solubilised proteins, ultracentrifugation of non-solubilised proteins or chromatography purification. Detergent concentration during solubilisation is kept within the CMC, which is the concertation at which detergent monomers assemble into larger non-covalently bound aggregates. Removal of the detergent is generally dependent on the CMC, detergents with higher CMC values can be dialysed against a larger volume over the span of days and with regular changing of the dialysis buffer. The gradual decrease of detergent concentration breaks the micelles down into dialysed monomers. For lower CMC detergents, hydrophobic beads are used for detergent removal followed by filtration or centrifugation of the detergent bound beads.

Alternatively, the co-translational insertion of proteins is a detergent-free method that utilises the CF protein expression systems (Nallani et al., 2011). With CF systems being open and modifiable, supplementation of expression mixture with synthetic cell membranes, including polymersomes, is possible. This insertion method relies on CF extract components for the transport of the newly translated peptide and insertion into the polymersome membrane. Although with a simpler incorporation scheme, there are downsides to this method; namely the low yields usually obtained with CF expression systems, the effect of synthetic membranes on CF expression machinery and protein yield, and the ability of the proteins to insert into the membranes. There are some attempts at utilising detergents within CF expression systems as an intermediary stage to aid in protein insertion into polymersome membranes.

#### CF proteopolymersomes

Currently, the most commonly used strategies for proteopolymersome production has been that of reconstitution following cell-based expression, however the use of cell-free expression and co-translational insertion has been gaining traction, with it successfully attempted on GPCRs, membrane structural proteins and other multimeric proteins.

Multiple different membrane proteins classes have been successfully inserted into polymersome surface, creating so called proteopolymersomes (Beales et al., 2017). Most of proteopolymersomes created have been using cell-based protein expression, followed by purification for post-translational protein insertion (Choi and Montemagno, 2005, Wang et al., 2013). However, CF systems have been used for simpler and quicker co-translational protein insertion. Preliminary testing showed the capability of co-translational protein insertion into planar polymer surfaces (Andreasson-Ochsner et al., 2012). Wheat germ based CF expression system supplemented with di- and triblock polymersomes was used for the expression and the co-translational insertion of GPCR dopamine receptor D2 (May et al., 2013) and the human structural protein claudin-2 (Nallani et al., 2011) into the polymersome membrane. The insertion of DRD2 was confirmed using flow cytometry, while ligand binding assays were used to confirm physiological folding and conformation of the receptor (May et al., 2013). Similarly for claudin-2, insertion was confirmed using Western blot analysis and surface Plasmon resonance (Nallani et al., 2011). Interestingly, the surface plasmon resonance results comparing claudin-2 containing polymersomes and liposomes showed increased anti-claudin-2 antibody interaction with polymersomes compared to liposomes, suggesting improved polymersome protein uptake and folding.

#### Purification strategies

Purification of proteopolymersomes is an important consideration for downstream applications. From protein assays to structural studies, purer protein samples are always highly desired. Polymersomes are highly suited for such applications due to their superior stability and hardiness against what most purification strategies demand. Common purification techniques such as centrifugation and filtration methods have been successfully utilised to separate polymersomes from unincorporated protein aggregates and the remainder of the protein expression components. Filtration membranes of selected pore sizes have been used for the retention of proteopolymersomes, while also washing away the smaller and undesirable impurities.

One key feature of polymersomes is their ease of labelling allowing for simple monitoring, identification and targeting of polymersomes. Labelling of the polymersome surface with target molecules that can be used to aid in their affinity purification, using targets such as biotin molecules (Lin et al., 2004). Size exclusion chromatography has also been used to separate proteopolymersome containing samples (Noor et al., 2012).

### Targeted polymersome insertion

Many proteins, which natively exist on membranes and those that do not, have been targeted for insertion and anchoring into polymersome membranes. Driven by functionalisation of polymersomes, special peptide sequences are being tested in their ability to target insertion into the polymersome as well as to link proteins to polymersomes that usually would not be. One of the main class of peptides that are being tested for this role are the antimicrobial peptides. These peptides, originally isolated from insect cells, target microbial membranes forming destabilising pores leading to cellular lysis. Studies have been done to analyse the efficiency these peptides target themselves into polymersome membrane. It’s been noted that the antimicrobial alamethicin peptide has no problem with targeting much thicker membranes of diblock polymersomes and is capable of inducing rupture (Vijayan et al., 2005). Another tested antimicrobial peptide, melittin, did show weak interactions with the polymersome, however vesicle rupture was not induced. Another antimicrobial peptide, cecropin A, was fused onto EGFP and tested as a membrane anchor for soluble proteins (Noor et al., 2012). The decoration of polymersomes with non-typical proteins will allow for increased understanding of membrane biophysics, the methods of protein insertion into membranes, as well as open up new methods for targeted polymersome delivery.

Non-antibacterial peptides have also been recently utilised for the anchoring of functional groups to polymersomes (Klermund et al., 2016). Three membrane binding peptides, the membrane binding region of C*b5*, the viral lysis protein L and the yeast syntaxin VAM3 were fused to EGFP and shown to reconstitute on triblock polymersomes, without affecting structural integrity. These peptide anchors have shown an alternative for surface decoration of polymersomes that is independent of polymer or lipid pre-conjugation.

### Future perspectives

Biomimetic membranes, such as liposomes, nanodiscs and polymersomes, have a significant role to play in the study of membrane proteins. Unlike native membranes, the usage of biomimetic alternatives allows for the simplification of protein reconstitution and incorporation methods, in addition to allowing the usage of membranes with novel and desirable characteristics. Polymersomes are very advantageous for such applications, the broad range of polymers available and possible conjugation allows for endless variations of polymersomes with a wide range of characteristics and functionalities, while the tight and rigid arrangements of polymers offer superior levels of stability compared to lipid based biomimetics. Focus on the *in vitro* insertion of proteins into the polymersome membrane using CF systems and purification of active proteopolymersomes will open up new avenues for downstream research and innovations. The development of microsomal monooxygenase proteopolymersome can offer a unique alternative for understanding of the function of microsomal monooxygenases, the cooperativity between them and their electron donors, the cooperativity between the different monooxygenases, as well as and elucidating its role in metabolism of its substrates. Answering these problems will aid in the understanding the relationship between drug and microsomal monooxygenases, how drugs can be designed with them in mind and how variations and side effects can be seen from person to person.

## Aims

The overall aim of this research work is to establish the suitability and the potential applications of polymersomes (polymer vesicles) as a biomimetic membrane environment for the solubilisation and stabilisation of membrane-bound proteins.

This will be tested out through the development of microsomal monooxygenase proteopolymersomes, in which functional microsomal monooxygenase is anchored to polymersomes. More specifically, consisting of the cytochrome P450 monooxygenase CYP3A4 and its two protein partners, the cytochrome P450 reductase and cytochrome *b5*, and will be used as a model system to verify the concept of proteopolymersome and to understand the role and functions of microsomal monooxygenases.

To achieve the overarching aim described above, the following specific objectives are planned:

* Using *in vitro* production methods, such as cell-free protein expression systems supplemented with polymersomes, for the co-translational insertion of these membrane-bound proteins into the biomimetic environment offered by the polymersomes.
* The purification of the (proteo)polymersomes.
* Testing the activity of the proteopolymersomes, as well as characterising the vesicle stability, shape, size and concentration.
* The elucidation of the roles of the different CYP3A4 peptide regions such as the N-terminus and the F-G helices and loops, in membrane binding, activity, and protein oligomerization.
* The development of peptide sequences capable of anchoring typically non-polymersome inserting proteins onto the polymersome surface in addition to usage as a tool for the surface functionalisation of polymersomes into proteopolymersomes.

**Chapter 2:**

# Methodology

2.1 MATERIALS

2.2 MOLECULAR CLONING

2.3 VESICLE PREPARATION

2.4 PROTEIN EXPRESSION

2.5 ANALYSIS OF PROTEIN EXPRESSION

2.6 PURIFICATION

2.7 CYP3A4 ACTIVITY ASSAY

2.8 CHARACTERISATION

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

CecA – Cecropin A

DPPE – 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine

DSPE – 1,2-distearoyl-sn-glycero-3-phosphoethanolamine

EGFP – Enhanced green fluorescent protein

GUV – Giant unilamellar vesicles

HLT – Polyhistidine tag-lymphotactin-TEV protease peptide

IPTG - Isopropyl β-D-1-thiogalactopyranoside

Mel – Melittin

PBD – Polybutadiene

PBS – Phosphate buffered saline

PCR – Polymerase chain reaction

PDMS - Poly(dimethylsiloxane)

PEG – Polyethelyne glycol

PMOXA - Poly(2-methyloxazoline)

PEO – Polyethelyne oxide

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

WT – Wild-type

List of Tables

**Table 2.1:** Primers used for the tagging and cloning of CYP3A4 into pET24a, pF25A, pT7CFE and pEU-E01 vectors.

**Table 2.2:** Primers used for the tagging and cloning of CPR into pEU-E01 vector.

**Table 2.3:** Primers used for the synthesis and cloning of Caeca, Mel and LAH4 peptides.

List of Figures

**Figure 2.1:** The protocol for vesicle formation using the film hydration and extrusion method.

**Figure 2.2:** Structural representation of lipid conjugated biotin, and its arrangement on polymersome membranes following rehydration.

**Figure 2.3:** Electroformation of liposomal GUVs.

## Materials

DNA polymerases*; Taq* and Q5 high-fidelity polymerases were obtained from New England Biolabs. Restriction enzymes; DpnI, NdeI, SgfI, PmeI, BamHI, NotI, EcoRI and EcoRV, and the T4 DNA ligase were obtained from New England Biolabs. QIAquick PCR purification kit, QIAprep spin miniprep kit and Aquic gel extraction kit were obtained from Qiagen. AxyPrep Maxiprep kit was obtained from Axygen. SeaKem LE Agarose was obtained from Lonza. NovaBlue (DE3) competent *E. coli* cells were obtained from Novagen.

The diblock Poly(butadiene-b-polyethylene oxide) polymer (subunits: PBD22-PEO13, Mw: 1200-b-600, PBD-PEO polymer), and the triblock poly(2-methyloxazoline-b-dimethylsiloxane-b-2-methyloxazoline) polymer (subunits: PMOXA20-PDMS54-PMOXA20, Mw: 1700-b-4000-b-1700) were obtained from Polymer Source Inc. (Canada). The 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-N-[biotinyl(polyethylene glycol)-2000] (DSPE-PEG-biotin, biotin conjugated to the lipid DSPE-PEG moiety), the 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC; 25 mg/ml) and 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine-N-(lissamine rhodamine B sulfonyl) (DPPE-rhodamine B, fluorescent rhodamine B conjugated to the lipid DPPE moiety) were from Avanti Polar Lipids, Inc (US). 200 nm membranes and filter supports were obtained from Avanti Polar Lipids.

Rosetta (DE3) Competent *E. coli* Cells were obtained from Novagen. *E. coli* S30 T7 High-Yield Protein Expression System and TnT T7 Insect Cell Extract Protein Expression System were obtained from Promega. 1-Step Human Coupled IVT Kit – DNA was obtained from ThermoFisher Scientific. Wheat germ cell-free expression system WEPRO7240 was from CellFree Sciences Co., Ltd. (Japan). Trace mineral supplements were obtained from ATCC. *δ*-Aminolevulinic acid hydrochloride was obtained from FORMEDIUM.

NuPAGE 10% Bis-Tris gels and Bolt 4-12% Bis-Tris Plus gels were obtained from Invitrogen. Coomassie Brilliant Blue - G-250 was obtained from Bio-Rad Laboratories. SYPRO Ruby Gel Stain and WesternBreeze Chemiluminescent Kit, anti-mouse and anti-rabbit were from ThermoFisher Scientific. Mouse rhodopsin antibody (1D4, sc-57432, 200 µg/mL) and rabbit His-probe antibody (H-15, sc-803, 100 µg/mL) were from Santa Cruz Biotechnology. AlexaFluor 488 goat anti-mouse (2 mg/ml) was obtained from ThermoFisher Scientific. IgG Gold (anti-Mouse) antibody was obtained from Sigma Aldrich.

SoftLink Soft Release Avidin Resin was obtained from Promega. 10x Phosphate Buffered Saline, Ultra-Pure Grade, pH 7.4 (diluted to a 1x solution of 137 mM sodium chloride, 2.7 mM potassium chloride and 10 mM phosphate buffer) was obtained from Vivantis. Vivaspin 500 (1000 kDa molecular weight cut-off/ pore size) was obtained from Sartorius.

The Vivid CYP3A4 Blue Screening Kit was from ThermoFisher Scientific. All other reagents were obtained from Sigma Aldrich.

## Molecular cloning

### CYP3A4

*CYP3A4 Nf14* (N-terminus trans-membrane domain deleted) gene sequence (appendix 1A) was kindly obtained from Professor Atkins’ lab (Gillam et al., 1993). PCR was used for the insertion of the membrane spanning N-terminus region to create the CYP3A4 WT sequence (appendix 1), as well as adding C-terminus polyhistidine (6x) tag and rhodopsin (1D4) tag. CYP3A4 WT and Nf14 constructs were cloned into pET24a vector for *E. coli* cell-based and CF expression, pF25A vector for insect CF expression, pT7CFE1 vector for HeLa CF expression and pEU-E01 (appendix 1B) for wheat germ CF expression.

#### pET24a-CYP3A4 cloning

NdeI and EcoRI restriction sites were used for *CYP3A4* insertion into the pET24a vector. Two unique forward primers were used for the WT and Nf14 constructs, a single reverse primer was utilised for the cloning of both constructs (Table 2.1). PCR was carried out using *Taq* polymerase, following standard molecular cloning techniques. PCR templates were digested with DpnI restriction enzyme and followed by PCR purification of amplified DNA. pET24a vector and CYP3A4 WT and Nf14 constructs were double digested with NdeI and EcoRI following suppliers protocol, followed by separation of constructs by 1% agarose gel electrophoresis. Gel extraction of gene constructs and linearised plasmid was carried out followed by vector and gene ligation using T4 DNA ligase following supplier’s protocols. Constructs were transformed into DH5α *E. coli* cells. Miniprep isolation of WT and Nf14 constructs (250 ng/µl) was carried out and sequencing was done to confirm sequence fidelity.

#### pF25A-CYP3A4 cloning

SgfI and PmeI restriction sites were used for *CYP3A4* insertion into pF25A vector. Two unique forward primers were used for the WT and Nf14 construct, a single reverse primer was utilised for the cloning of both constructs (Table 2.1). PCR was carried out using *Taq* polymerase, followed by template digestion with DpnI restriction enzyme and PCR purification of amplified gene product. pF25A vector and CYP3A4 WT and Nf14 constructs were double digested with SgfI and PmeI followed by separation of constructs and linearised plasmid by 1% agarose gel electrophoresis and the gel extraction of constructs using QIAquick gel extraction kit. T4 ligation of vector and gene was carried out and the constructs were transformed into DH5α *E. coli* cells. Miniprep isolation of WT and Nf14 constructs (250 ng/µl) was carried out and sequencing was done to confirm sequence fidelity.

|  |  |  |
| --- | --- | --- |
| **Table 2.1:** Primers used for the tagging and cloning of *CYP3A4* into pET24a, pF25A, pT7CFE and pEU-E01 vectors. | | |
| **pET24a** | **Fwd (Nf14)** | **5’** T ATA CAT ATG ATG GCT CTG TTA TTA GCA GTT TCC CTG GTG CTC CTC **3’** |
| **Fwd (WT)** | **5’** T ATA CAT ATG ATG GCT CTG ATT CCG GAT CTG GCG ATG GAA ACC TGG CTG TTA TTA GCA GTT TCC CTG GTG CTC CTC **3’** |
| **Rev** | **5’** G CTC GAA TTC TTA ATG ATG ATG ATG ATG ATG GGT CGA CGC TCC **3’** |
| **pF25A** | **Fwd (Nf14)** | **5’** TAA AGC GAT CGC ATG GCT CTG TTA TTA GCA GTT TCC CTG GTG CTC CTC **3’** |
| **Fwd (WT)** | **5’** TAAAGC GAT CGC ATG GCT CTG ATC CCC GAC CTG GCT ATG GAG ACC TGG CTG TTA TTA GCA GTT TCC CTG GTG CTC CTC **3’** |
| **Rev** | **5’** ATT CGT TTA AAC TTA ATG ATG ATG ATG ATG ATG GGT CGA CGC TCC **3’** |
| **pT7CFE1**  **and**  **pEU-E01** | **Fwd (WT and Nf14)** | **5’** AA AGA TAT CAA AAT AAA ATA TAA AGC GAT CGC ATG GCT CTG **3’** |
| **Rev** | **5’** TTT TGC GGC CGC TTT TTA TTA GGC GGG GGC CAC CTG GGA GGT CTC GGT ATG ATG ATG ATG ATG ATG GGT CGA CGC **3’** |

#### pT7CFE1-CYP3A4 and pEU-E01-CYP3A4 cloning

EcoRv and NotI restriction sites were used for *CYP3A4* insertion. One forward primer and one reverse primer were utilised for the cloning of both WT and Nf14 constructs in both pT7CFE1 and pEU-E01 vectors (Table 2.1) from pF25A-CYP3A4 WT/Nf14 vectors. PCR was carried out using Q5 high-fidelity polymerase using manufacturer’s protocol, followed by template digestion with DpnI restriction enzyme and PCR purification of amplified genes. pT7CFE1 and pEU-E01 vectors, and CYP3A4 WT and Nf14 constructs were double digested with EcoRv and NotI followed by separation of constructs by 1% agarose gel electrophoresis and gel extraction. T4 ligation of vectors and genes was carried out and the constructs were transformed into NovaBlue (DE3) *E. coli* competent cells. Miniprep isolation of WT and Nf14 constructs (250 ng/µl) was carried out and sequencing was done to confirm sequence fidelity. Midiprep of constructs (1µg/µl) was carried out to obtain high concentration and purity.

### CPR

Untagged WT *CPR* gene was obtained from Origene, USA (Cat # SC100401). PCR, using Q5 high-fidelity polymerase, was carried out to insert a C-terminus 1D4 tag, and EcoRV and NotI restriction sites at the N- and C- terminus respectively (Appendix 2A; primers shown in Table 2.2). The gene construct was cloned into pEU-E01 vector for wheat germ CF expression (Appendix 2B). The DNA template was digested with DpnI and the amplified PCR product was purified. pEU-E01 and purified PCR product were double digested with EcoRV and NotI, and separated using 1% agarose gel electrophoresis. T4 ligation of linearised pEU-E01 and CPR was carried out and the construct was transformed into NovaBlue (DE3) *E. coli* competent cells. Miniprep isolation of pEU-E01-CPR construct (250 ng/µl) was carried out and sequencing was done to confirm sequence fidelity. Midiprep (1 µg/µl) of the construct was carried out to obtain high concentration and purity.

|  |  |  |
| --- | --- | --- |
| **Table 2.2:** Primers used for the tagging and cloning of *CPR* into pEU-E01 vector. | | |
| **pEU-E01** | **Fwd** | **5’** GCA GCG ATA TCT AAC GCA TGA TCA ACA TGG GAG ACT CCC ACG **3’** |
| **Rev** | **5’** CTA TTA GGC GGG GGC CAC CTG GGA GGT CTC GGT GCT CCA CAC GTC CAG GG **3’** |

### C*b5*

Synthesised C-terminus 1D4 tagged *Cb5* gene was obtained from Integrated DNA Technologies (Singapore) (appendix 3A, 3B). The gene construct, containing EcoRV and NotI restriction sites at the N- and C-terminus respectively was double digested and separated using 1% agarose gel electrophoresis. T4 ligation of linearised pEU-E01 vector and digested *Cb5* was carried out and the construct was transformed into NovaBlue (DE3) *E. coli* competent cells. Miniprep isolation of pEU-E01-C*b5* construct (250 ng/µl) was carried out and sequencing was done to confirm sequence fidelity. Midiprep of the construct (1 µg/µl) was carried out to obtain high concentration and purity.

### Membrane peptides

The antimicrobial peptides cecropin A (AA sequence: KWKLFKKIEKVGQNIRDGIIKA-GPAVAVVGQATQIAK), melittin (AA sequence: GIGAVLKVLTTGLPALISWIKRKRQQ) and LAH4 (AA sequence: KKALLALALHHLAHLALHLALALKKA) were fused to eukaryotic green fluorescent protein (EGFP) with a polyalanine linker, and cloned into the pET24a vector for *E. coli* cell-based and CF expression and into pEU-E01 vector for wheat germ CF expression.

pET24a-HLT-EGFP construct containing a polyhistidine tag-lymphotactin-TEV protease (HLT; AA sequence: HHHHHHSGAFEFKLPDIGEGIHEGEIVKWFVKPGDEVNEDDVLC EVQNDKAVVEIPSPVKGKVLEILVPEGTVATVGQTLITLDAPGYENMTTGSDTGENLYFQGGSMGS) peptide sequence fused upstream to EGFP was obtained. PCR was carried out for the synthesis of the peptides (table 2.3). The PCR peptide products, flanked by NdeI and BamHI restriction sites were double digested, and the pET24a-HLT-EGFP plasmid was similarly double digested and linearised to remove the HLT peptide. The peptide sequences and linearised plasmid were separated using 1% agarose gel electrophoresis. T4 ligation of pET24a-*x*-EGFP and peptide sequences was carried out and the constructs were transformed into NovaBlue (DE3) *E. coli* competent cells. Miniprep isolation of pET24a-CecA/Mel/LAH4-EGFP constructs was carried out and sequencing was done to confirm sequence fidelity. C-terminus polyhistidine (6x) tag was added to all three constructs (primers in table 2.3), using standard molecular cloning techniques.

Mel-EGFP and HLT-EGFP constructs were cloned into pEU-E01 vector for wheat germ CF expression. PCR was used to insert EcoRV and HindIII restriction sites at the N- and C-terminus respectively, as well as a C-terminus polyhistidine (6x) tag for Mel-EGFP (primers shown in table 2.3). Amplified PCR product and pEU-E01 vector were double digested with EcoRV and HindIII, and separated using 1% agarose gel electrophoresis. Following gel extraction of gene sequences and linearised plasmid, T4 ligation was carried out and the constructs were transformed into NovaBlue (DE3) *E. coli* competent cells. Miniprep isolation of pEU-E01-Mel/HLT-EGFP constructs was carried out and sequencing was done to confirm sequence fidelity.

|  |  |  |
| --- | --- | --- |
| **Table 2.3:** Primers used for the synthesis and cloning of CecA, Melittin and LAH4 peptides. | | |
| **CecA**  **pET24a** | **Fwd** | **5’** TAT ACA TAT GAA GTG GAA ACT TTT CAA AAA AAT CGA AAA AGT GGG TCA GAA TAT CCG TGA TGG TAT CAT TAA GGC GGG GCC AGC CGT AGC **3’** |
| **Rev** | **5’** ATT CGG ATC CTG CGG CGG CCG CTG CGG CCG CTG CCG CTG CTT TGG CAA TCT GCG TCG CAC CGC CTA CAA CCG CTA CGG CTG GCC CCG CCT TA **3’** |
| **Melittin**  **pET24a** | **Fwd** | **5’** TAT ACA TAT GGG GAT TGG GGC CGT TCT GAA AGT CCT GAC CAC TGG CCT GCC CGC CCT CAT TAG CTG GAT TAA ACG CAA GCG CCA GCA GG **3’** |
| **Rev** | **5’** ATT CGG ATC CCC GCG GCC GCA GCC GCT GCA GCC GCG GCT GCG GCG GCA GCT GCA GCT GCC GCA GCA GCA GCC TGC TGG CGC TTG CGT TTA A **3’** |
| **LAH4**  **pET24a** | **Fwd** | **5’** TAT ACA TAT GAA GAA AGC ACT GTT GGC CCT GGC ACT TCA CCA TCT GGC GCA CCT GGC ACT CCA CCT GGC GTT GGC CCT TAA AAA AGC AG **3’** |
| **Rev** | **5’** ATT CGG ATC CCG CGG CCG CCG CCG CCG CCG CTG CAG CTG CTG CCG CTG CGG CTG CTG CCG CGG CCG CAG CTG CTT TTT TAA GGG CCA ACG **3’** |
| **Addition of polyhistidine (6x) tag** | **Fwd** | **5’** TAATACGACTCACTATAG **3’** |
| **Rev** | **5’** ACG GAG CTC GAA TTC TTA ATG ATG ATG ATG ATG ATG CTT GTA CAG CTC GTC CAT **3’** |
| **HLT-EGFP**  **pEU-E01** | **Fwd** | **5’** CAC CAA GAT ATC ATA ACT TAT GCA CCA CCA CC **3’** |
| **Rev** | **5’** GTT ATG CAA GCT TCT TTT TTA CTT GTA CAG CTC G **3’** |
| **Mel-EGFP**  **pEU-E01** | **Fwd** | **5’** CAC CAA GAT ATC ATA ACT TAT GGG GAT TGG GGC C **3’** |
| **Rev** | **5’** GTT ATG CAA GCT TCT TTT TAG GCG GGG GCC ACC TGG GAG GTC TCG GTA TGA TGA TGA TGA TGA TGC TTG **3’** |

## Vesicle preparation

Polymersomes and liposomes were created using the film hydration and extrusion method (figure 2.1) (Nallani et al., 2011). Giant unilamellar liposomes were created using electroformation method for the reconstitution of *E. coli* expressed and polyhistidine purified proteins.

### Polymersomes

Diblock PBD-PEO (Polybutadiene-Polyethylene oxide) polymers were dissolved in chloroform at 100 mg/ml concentration. The polymer solution was added to a test tube and dried using nitrogen gas aeration, followed by further drying in a vacuum. For the creation of vesicles containing membrane-anchored biotin, DSPE-PEG-Biotin (1,2-distearoyl-sn-glycero-3-phosphoethanolamine-Polyethelyne glycol-Biotin) was added and mixed to the polymer film at 5% molar equivalent mix (figure 2.2). Similarly, for rhodamine fluorescent vesicles, DPPE-lissamine-Rhodamine B was added and mixed to the polymer film at 1% molar equivalent. The polymer film was then rehydrated in a 1x PBS solution at final concentration of 10 mg/ml of PBD-PEO polymer. Following rehydration, extrusion was carried out using 200 nm filter and vesicles were dialysed in 1x PBS solution overnight.

|  |
| --- |
| Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-05-03 at 15.46.52.png |
| **Figure 2.1:** The protocol for vesicle formation using the film hydration and extrusion method. |

|  |
| --- |
|  |
| **Figure 2.2:** Structural representation of lipid conjugated biotin, and its arrangement on polymersome membranes following rehydration. This structural arrangement is similarly observed with lipid conjugated rhodamine B on the membranes of polymersomes. |

### Liposomes

#### Extrusion method

Liposomes were made using POPC lipids at 1 mg/ml concentration, similarly rehydrated in 1x PBS solution. For the creation of vesicles containing membrane-anchored biotin, DSPE-PEG-Biotin was added to the film at 5% molar equivalent mix. Similarly, for rhodamine fluorescent vesicles, DPPE-lissamine Rhodamine B was added to the film at 1% molar equivalent. Following rehydration, extrusion was carried out using 200 nm filter membranes and vesicles were dialysed in 1x PBS solution overnight.

#### Electroformation method

Liposomal GUVs were created using electroformation method (figure 2.3). Briefly, POPC lipid film (with 1% molar equivalent DPPE-lissamine Rhodamine B) was dried on a glass slide coated with indium tin oxide and covered with an additional protein film of the desired membrane protein. A 40 mM sucrose solution was sealed onto the lipid/protein film and a current run through it.

|  |  |
| --- | --- |
| **1.** Lipid layer is added to ITO-coated glass slide and left to dry for one hour. Protein layer on top and left to dry for 6 hours. | Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-05-02 at 15.25.48.png |
| **2.** O-ring added onto protein-lipid layer and sealed with grease. 40 mM Sucrose solution is added to form GUV environment. | Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-05-02 at 15.18.27.png |
| **3.** Glass slide is added to seal the protein-lipid and sucrose solution. A current is run through the sealed system to start vesicle formation. | Macintosh HD:Users:HossamMacbook:Desktop:Screen Shot 2017-05-02 at 15.18.27.png |
| **Figure 2.3:** Electroformation of liposomal GUVs. | |

## Protein expression

### *E. coli*

Cell-based protein expression was carried out using C41 (DE3) or Rosetta *E. coli* cells. Cells were transformed with pET24a vector constructs using CaCl2 heat-shock method, and inoculation cultures were carried out in LB media. Protein expression was done in terrific broth media using a final 1 mM IPTG concentration or with auto-induction recipe (Studier, 2005). All CYP3A4 expression media was additionally supplemented with 1 mM trace elements and 1 mM *δ-*aminolevulinic acid. Expression was done at 30–37oC for 6 – 24 hours. Cells were harvested post-expression and centrifuged at 5000 rpm for 30 minutes at 4oC. Cell pellets were either analysed for protein expression or stored at –20°C.

S30 T7 high-yield protein expression system (Promega) was used for *E. coli* CF protein expression. Reactions were prepared following manufacturers protocol. Polymersome supplementation of CF reaction mixture was done using 1.5 – 2.5 μl of 10 mg/ml PBD-PEO polymersomes per 50 μl CF reaction. CYP3A4 expression media was additionally supplemented with 1 mM trace elements and 1 mM *δ-*Aminolevulinic acid (heme precursor). Expression was carried out at 37°C for 1-6 hours, and 30°C for 24 hours. Completed reactions were either analysed for expression or aliquoted and stored at –20°C.

### Insect CF

Coupled transcription and translation CF expression was carried out using TnT T7 insect cell extract protein expression system. CF reactions were prepared following manufacturers protocol. Polymersome supplementation of CF reaction mixture was done using 1.5 – 2.5 μl of 10 mg/ml PBD-PEO polymersomes per 50 μl CF reaction. CYP3A4 expression media was additionally supplemented with 1 mM trace elements and 1 mM *δ-*Aminolevulinic acid. Expression was carried out at 30°C for 6 hours. Completed reactions were either analysed for expression or aliquoted and stored at –20°C.

### HeLa CF

1-Step Human Coupled IVT Kit was used for coupled transcription and translation HeLa CF protein expression. Reactions were prepared following manufacturers protocol. Polymersome supplementation of CF reaction mixture was done using 3 μl of 10 mg/ml PBD-PEO polymersomes per 25 μl CF reaction. Cofactor supplementation of the expression mixture was carried out with 5 µM hemin, 5 µM FAD and 5 µM FMN. Expression was carried out at 30°C for 6 hours. Completed reactions were either analysed for expression or aliquoted and stored at –20°C.

### Wheat germ CF

Wheat-germ based CF expression was carried out using the linked WEPRO 7240 expression system, at both small scale (10 µl WEPRO wheat germ extract) and medium scale (50 µl WEPRO wheat germ extract). Transcription was carried out for each protein independently for 6 hours at 37oC using supplied transcription mix and the midiprep of the constructs used at a concentration of 1µg/µl. This was followed by co-enzyme translation using the generated mRNA, as well as the supplied WEPRO and SUBA-MIX in a biphasic translation system. The lower phase during translation is composed of WEPRO, generated mRNA and vesicles/water in a 1:1:1 ratio, while the mRNA ratio itself is 1:2:2 ratio of CYP3A4:CPR:C*b5*/water. The upper phase is SUBA-MIX supplemented with 5 µM hemin, 5 µM FAD and 5 µM FMN. Translation was carried out for 20 hours at 16oC. Completed reactions were either analysed for expression and activity or aliquoted and stored at 4°C.

## Analysis of protein expression

### SDS-PAGE

Protein gel electrophoresis was carried out on pre-cast NuPAGE 10% Bis-Tris and Bolt 4-12% Bis-Tris Plus gels. 10 µl of protein sample were mixed with 10 µl of 2x loading dye and loaded per well, unless stated otherwise. The protein samples with LDS loading buffer and NuPAGE reducing agent, mixed for a final 1x concentration were heated at 70oC for 10 minutes. Samples were loaded onto gels and run according to manufacturer’s protocols. Staining was carried out using Coomassie blue G250 dye (Bio-Rad) or SYPRO Ruby protein gel stain, following manufacturers’ recommended protocol. Imaging was done using ImageQuant LAS 500 (GE Healthcare Life Sciences). Protein band quantification and analysis was carried out using QuantityOne software (Bio-Rad Laboratories).

### Western blot

Protein transfer was done using iBlot 2 (ThermoFisher Scientific) dry blotting method. Western blot was done using WesternBreeze Chemiluminescent Kit and performed according to manufacturer’s manual. Proteins were tagged using either mouse anti-rhodopsin antibody or anti-histidine tag antibodies. Imaging was done using ImageQuant LAS 500 (GE Healthcare Life Sciences).

## Purification

### Polyhistidine tag-Nickel purification

Purification of histidine tagged proteins was carried on using HisTrap HP 5 ml column (GE Healthcare Life Sciences) on an ÄKTA FPLC system (GE Healthcare Life Sciences). Cell pellets were lysed using sonication (10 seconds on, 20 seconds off for total of 2 minutes on; 70% amplitude). Lysis buffer was composed of 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 1 mM PMSF. Running buffer is composed of 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 50 mM imidazole, the elution buffer is composed of 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 500 mM imidazole. Purified protein samples were collected, aliquoted and analysed for purification or stored at -20°C in 25% glycerol solution.

### Biotin-avidin purification

Avidin purification of biotin-tagged vesicles was carried out using Softlink Soft Release Avidin Resin. The avidin resin was incubated with the expression mixture in a 1:3 ratio for 20 hours at 4oC with shaking at 1000 rpm. The resin was washed three times using 1x PBS solution at 4 bed volumes. The vesicles were eluted off the resin using 10 mM biotin solution in 1x PBS at 1:3 ratio of elution buffer to resin for 3 hours at 4oC with shaking at 1000 rpm.

### Filtration

Vesicles were purified from the CF reaction mixtures using 1000 kDa molecular weight cut-off Vivaspin 500 columns. The columns were pre-washed with 200 µl of filtered 1x PBS and centrifuged (6000 rpm) at 4°C for 20 minutes. The expression mixture was loaded onto the Vivaspin column and centrifuged (6000 rpm) at 4°C for 30 minutes. Three subsequent washes with 1x PBS were carried out with centrifugation (6000 rpm) at 4°C for 30 minutes. The retentate was either analysed or stored at 4oC.

### Centrifugation

1 ml of protein or vesicle containing samples were centrifuged at 10 000 or 14 500 rpm for 30 minutes, 1 hour or 2 hours at 4oC for the pelleting of protein aggregates and CF expression components. The pellet and the supernatant fractions were separated. The pellet was resuspended in 250 µl of 1x PBS solution. The samples were either analysed or stored at 4oC.

### Size exclusion chromatography

Size exclusion chromatography of protein and vesicle samples was carried out using Superdex 200 5/150 GL column (GE Healthcare Life Sciences) on an ÄKTA FPLC system. 100 µl fractions of the samples were collected and analysed on SDS-PAGE or Western blot, and stored at 4oC.

## CYP3A4 activity assay

Activity of the CYP3A4 monooxygenase system was tested using Vivid CYP3A4 Blue Screening Kit. CYP3A4 Baculosomes (ThermoFisher Scientific) activity was monitored using manufacturers protocol. Prepared CYP3A4 monooxygenases preparations were tested in similar protocol; 49.5 µl of CYP3A4 preparation was supplemented with 0.5 µl of Vivid Regeneration System and left to equilibrate for 20 minutes. Activity was started with 10 µl start mixture composed of 0.5 µl of 10 µM BOMCC substrate, 0.3 µl of 10 mM NADP+ and 9.2 µl of Vivid reaction buffer. BOMCC is a blocked dye with two oxidation sites, the oxidation of which by CYP3A4 releases a highly fluorescent dye. The enzyme activity kinetics was monitored regularly for 2 hours with excitation at 415 nm and emission at 460 nm. Controls were set-up for the activity analysis, including negative control (CF expression product when no mRNA is added during the expression stage), and positive control (use of CYP3A4 containing Baculosomes supplied with the Vivid Blue CYP3A4 assay kit).

## Characterisation

### Dynamic light scattering

Zetasizer Nano ZS (Malvern Instruments) was used for analysis of vesicle size, polydispersity index and vesicle count via dynamic light scattering. DLS measures the size distribution of small particles in solution (PBS). When the light hits the polymersome and liposome vesicles, the scattered light travels in differing directions with intensity fluctuation. These fluctuations can be correlated to deduce the average size of measured vesicles and their distribution.

300 µl of 10-fold diluted samples of either freshly prepared vesicles (Methodology 2.3), biotin affinity purified polymersomes samples (Methodology 2.6.2) or centrifuge (10,000 or 14,500 rpm for 30 minutes, 1 hour or 2 hours) purified vesicle samples (Methodology 2.6.4) were analysed using DLS in triplicates (of ten individual readings each time), with the mean of the intensity readings (total of 30 readings) plotted against vesicle size. All DLS analysis of the vesicles were carried out in PBS buffer, at room temperature (25oC) with readings taken at 173o backscatter.

### Flow cytometry

Biotin and rhodamine tagged (proteo)polymersomes were used for flow cytometry analysis. Following wheat germ CF protein expression, the (proteo)polymersomes were incubated and tethered to the avidin resin. The (proteo)polymersome-resin mix was washed with 1x PBS and incubated with anti-1D4 antibody for 2 hours. Unbound primary antibodies washed away with 1x PBS and the proteopolymersome-resin incubated with secondary AlexaFluor488 antibody. Unbound secondary antibodies were washed with 1x PBS and the proteopolymersomes were eluted off the avidin resin using 10 mM Biotin. The proteopolymersome elutant was run through the flow cytometry (BD LSR II), and the populations of polymersomes was tracked using yellow-green laser (ex: 561 nm, em: 582), protein-antibody complex was tracked using blue laser (ex: 488 nm, em: 530 nm).

### Fluorescence microscopy

Fluorescence microscopy was carried out on EGFP or rhodamine B tagged vesicles using an Olympus BX51 microscope equipped with FITC (EGFP) and (PE-A) rhodamine filters. For detection of EGFP and fluorescent AlexaFluor 488 antibodies, the samples were excited at 485 nm and emission detected at 520 nm. For detection of rhodamine, the samples were excited at 550 nm, and the emission detected at 590 nm.

### Transmission electron microscopy

Avidin purified (proteo)polymersome samples were imaged using transmission electron microscopy (TEM) on a Libra 120 Zeiss. Immunostaining with mouse anti-1D4, and anti-mouse gold nanoparticle conjugated antibodies was carried out while (proteo)polymersomes were attached to avidin beads, with triplicate 1x PBS wash after every stage. Samples were prepared by direct application to plasma-treated copper grids followed by negative staining with 4% osmium tetroxide.

**Chapter 3:**

# Cloning and Expression of CYP3A4

3.1 INTRODUCTION

3.2 AIMS

3.3 EXPERIMENTATION

3.4 CLONING

3.5 EXPRESSION

3.6 DISCUSSION

3.7 SUMMARY

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

EMCV – Encephalomyocarditis virus

IRES – Internal ribosome entry site

PBD – Polybutadiene

PEO – Polyethelyne oxide

UTR – Untranslated region

WT – Wild-type

List of Figures

**Figure 3.1:** SDS-PAGE of CYP3A4 expression in C41 (DE3) *E. coli*, and harvested cell pellets.

**Figure 3.2:** Western blots of CYP3A4 protein expression in insect, HeLa and wheat germ CF systems.

## Introduction

Typically for the expression of human proteins, such as the CYP3A4 enzyme, cell-based systems have been utilised. *E. coli* by far being the most commonly utilised protein expression system, even for proteins of eukaryotic origin. However, with microsomal monooxygenases, while *E. coli* cells have been previously used, they are not an ideal protein expression system as it does not offer a sufficient yield of expressed protein, and usually what is expressed is found accumulated within inclusion bodies as non-functioning aggregates. Additionally, there is still the problem of reconstituting these aggregates onto membranes to create a functional monooxygenase system. Alternatively, eukaryotic protein expression systems, such as human or insect cells, have been utilised for the expression and targeted delivery of CYP3A4 into the membrane of the endoplasmic reticulum followed by the subsequent purification of microsomes containing folded and active proteins. These eukaryotic microsomes, although functional, were often not as pure as desired and lacked control of contents (*e.g.* which enzymes were present and at what concentration).

CF protein expression systems, similar in characteristics to their cell-based origins, are similarly expected to have varying degrees of expression yields of the human CYP3A4. The prokaryotic *E. coli* CF system, usually a high yield system, can struggle with expression of eukaryotic proteins, although if successful, the high yields and relatively low cost of the system will make this a highly desirable expression system (Kigawa et al., 2004). Eukaryotic CF systems, in this case insect, wheat germ and HeLa based, are more capable of expressing human CYP3A4, however the yields are expected to vary between them, with wheat germ CF yielding much higher levels of expressed protein in comparison to insect or HeLa systems. Wheat germ CF does not have as much post-translational capabilities as the other two systems, although its capabilities should be sufficient for expression of functional heme-containing membrane proteins.

The presence of biomimetic membranes within CF expression mixtures has shown to have an effect in the yields of expressed proteins obtained, possibly favourably through molecular crowding or stabilisation properties, or possibly detrimentally through interference and inhibition of mobility. Liposome supplemented *E. coli* CF expression system has shown that it can have an overall positive effect on expression yields (Bui et al., 2008). The analysis of EGFP expression using CF expression mixture showed increase in both mRNA levels (showing enhanced transcription), and fluorescence levels (indicating enhanced protein translation). It is thought that this enhancement is due to the stabilising effect of the liposomal membrane on T7 RNA polymerase, as well as providing the ribosomes with scaffold for the synthesis of polypeptide chains. This enhancing effect however was only observed for zwitterionic, neutral liposome vesicles, meanwhile charged liposomes supplemented CF systems lead to a decrease in fluorescence levels, perhaps with the charge affecting ribosomal translation and protein folding. In the analysis of polymersome supplementation effect on CF it’s been shown that yield is only affected after a certain point (0.5 µg/µl), at which protein yield levels in a wheat germ based CF system decreased (Zapf et al., 2015). Ideally when using membrane supplemented CF systems is to have as high membrane presence as possible, to maximise chance interactions between peptide and polymersome, while also not affecting transcription or translation levels.

## Aims

The cloning of the WT and the mutant (Nf14, trans-membrane region deficient mutant) *CYP3A4* genes into four unique and specialised expression vectors; pET24a, pF25A, pT7CFE1 and pEU-E01. Their expression in (DE3) *E. coli* cells followed by the expression of WT and mutant CYP3A4 in prokaryotic and eukaryotic based CF protein expression systems to identify the CF system capable of expressing human microsomal monooxygenases with highest yields and low relative cost. To observe the effect of diblock copolymer composed polymersome presence within the protein expression mixture on the expression yields in the CF systems.

## Experimentation

### Cloning

Standard molecular biology techniques of PCR, DNA digestion, ligation, transformation and isolation were used for the cloning of both WT and Nf14 *CYP3A4* genes into pET24a, pF25A, pT7CFE1 and pEU-E01 vectors, downstream of the expression promoter in the multiple cloning sites, as described in the Methodology section 2.2. Polyhistidine (6x) and rhodopsin (1D4) tags were cloned at the c-terminus.

### Expression

Transformation of the pET24a-CYP3A4 WT and Nf14 plasmids into C41 *E. coli* cells, and grown in terrific broth auto-induction media. Expression was carried out at 30oC, and monitored at 24 and 48 hours, as described in the Methodology section 2.4. Collected cell samples were spun down, and resuspended in a LDS sample loading buffer and NuPAGE reducing agent buffer mix. Samples were run through an SDS-PAGE and stained with Coomassie blue G250 dye.

CF protein expression was carried out in *E. coli,* insect, HeLa and wheat germ based expression systems, following manufacturers’ protocols. The expression systems chosen have been shown to express similar proteins to CYP3A4, including other cytochrome P450 enzymes, as well as under similar conditions. CF reactions were supplemented with polymersome vesicles; 0.3 and 0.5 µg/µl total PBD-PEO polymer concentration in *E. coli* and insect CF systems, 0.6 µg/µl total PBD-PEO polymer concentration in HeLa CF system, and 0.3 µg/µl total PBD-PEO polymer concentration in wheat germ CF. Following expression, polymersome deficient and replete samples were run through an SDS-PAGE followed by Western blot detection (polyhistidine and 1D4 tag antibody labelling).

## Cloning

WT and Nf14 *CYP3A4* was successfully cloned into the multiple cloning site on the pET24a, pF25A, pT7CFE1 and pEU-E01 vectors. DNA sequencing was carried out confirming the desired sequence fidelity. Plasmid isolation, via mini- and midi-prep, obtained vectors at high concentration (300 – 1000 ng/µl) and high purity.

## Expression

CYP3A4 was successfully expressed in C41 (DE3) *E. coli* cells (figure 3.1a). Over-expression protein bands visible among presence of other native *E. coli* proteins after 24 and 48 hours of expression in terrific broth auto-induction media. No observable difference in yields of WT or Nf14 CYP3A4 protein levels was seen. Also, no major difference was observed in yield of CYP3A4 after 48 hours compared to 24 hours, however, there was increased levels of non-target proteins (at lower kDa values). Collected cell pellet samples at 24 hours showed red coloration, characteristic of heme protein expression, for WT and Nf14 CYP3A4 protein expression, and was not observed for empty pET24a vector expression (figure 3.1b).

|  |
| --- |
| **a)** |
| **b)** |
| **Figure 3.1: a)** SDS-PAGE of WT (lanes 1 and 2) and Nf14 mutant (lanes 3 and 4) CYP3A4 expression in C41 (DE3) *E. coli* cells, at 24 hours (lanes 1 and 3) and 48 hours (lanes 2 and 4). **b)** Collected cell pellets of C41 (DE3) *E. coli* cells at 24 hours post-induction for pET24a (1), pET24a-CYP3A4 Nf14 (2), and pET24a-CYP3A4 WT (3) |

The Western blot of the CYP3A4 WT and Nf14 *E. coli* CF expression showed no protein detection bands, indicating the protein is not expressed by the S30 *E. coli* CF system. Similarly, CYP3A4 Nf14 expression was not detected in the Insect CF expression system (with or without polymersome supplementation). However, the CYP3A4 WT protein expression was detected, showing slight increase in protein yields with polymersome supplementation (figure 3.2a).

|  |
| --- |
| **a)** |
|  |
| **b)** |
|  |
| **c)** |
| **Figure 3.2: a)** Western blot of Nf14 (lane 1) and WT (lane 2-4) CYP3A4 expression in insect CF expression system. Polymersomes supplemented insect CF expression system for WT CYP3A4 with 0.3 (lane 3) and 0.5 (lane 4) µg/µl of total PBD-PEO polymer concentration in the expression mixture. **b)** Western blot of Nf14 (lane 1 and 2) and WT (lane 3 and 4) CYP3A4 expression in HeLa CF expression system. Polymersomes supplemented HeLa CF expression system for CYP3A4 expression with 0.6 µg/µl total PBD-PEO polymer concentration in the expression mixture (lanes 2 and 4). **C)** Western blot of Nf14 (lane 1 and 2) and WT (lane 3 and 4) CYP3A4 expression in wheat germ CF expression system. Polymersomes supplemented wheat germ CF expression system for CYP3A4 expression with 0.3 µg/µl total PBD-PEO polymer concentration in the expression mixture (lanes 2 and 4). |

CYP3A4 Nf14 expression was detected in the Western blots of the HeLa and wheat germ CF expression systems (figure 3.2 b and c) albeit at lower levels than observed for the WT constructs. It’s also been observed that the presence of polymersomes within the CF expression mixture resulted in a slight increase in the protein yields for both the WT and Nf14 CYP3A4 in the HeLa and wheat germ based systems.

## Discussion

The cloning of CYP3A4 WT and Nf14 into the specialised expression vectors is important for optimisation of protein yields. The pET24a vector used for *E. coli* based expression utilises the standard T7 expression system and while it has some benefits; *e.g.* the T7 constitutive promoter region for controlled expression of recombinant proteins, it has no specialised “factor sequences” that aid in the expression of complex eukaryotic proteins such as human CYP3A4. While that vector was sufficient to drive CYP3A4 expression in C41 *E. coli* cells, it provided no aid to the S30 *E. coli* CF system, which was unable to transcribe and translate the protein. Factors inhibiting CYP3A4 expression in *E. coli* CF could be occurring at the translation stage *e.g.* from inefficient codon usage. With the redundancy in the genetic code, most species have codons that are more likely to be recognised and undergo transcription due to variations in their tRNA levels. Having high levels of codons in an mRNA sequence that have low frequency in *E. coli* is detrimental to protein expression, acting as a bottleneck to translation. Generally, when expression of a difficult protein is desired in an S30 CF system, additional enhancing elements are required. Both 5’ and 3’ UTR sequences have been used to inhibit exonuclease mediated degradation and inhibition of any translational inhibitors in S30 CF extract (Ahn et al., 2005, Son et al., 2006). N- and C-terminus fusion tags such as CAT and human immunoglobulin κ light chain constant domain have shown large improvements in S30 yields of proteins typically not expressed proteins through enhancement of translation initiation, polypeptide stabilisation and inhibition of degradation (Palmer et al., 2006), while alteration of the mRNA structure was also shown to improve stability (Ahn et al., 2008) and the initiation of translation (Paulus et al., 2004).

As expected, the C41 *E. coli* cells were able to express CYP3A4 in high levels (figure 3.1a), with no significant increase in yields observed between 24 and 48 hours, however there is an increase in degradation products at 48 hours noted with the thicker bands at lower kDa values. The characteristic reddish colour to the cell pellets indicates the capabilities of *E. coli* cells to convert *δ-*aminolevulinic acid to the heme cofactor (Gillam, 2007), and have it presented to the enzyme while in a form capable of folding and accepting the heme functional group. However, to utilise the expressed protein in further downstream processes, such as reconstitution into biomimetic vesicles, it would require time-consuming solubilisation procedures, purification from the native *E. coli* proteins and reconstitution.

The vectors used for the insect, HeLa and wheat germ CF systems are more specialised for eukaryotic protein expression and aid in optimising expression yields. The pF25A vector of the insect CF expression contains a 5’ and a 3’ polyhedrin UTRs and an additional 3’ poly(A) sequence downstream from it. The polyhedrin sequence is a well utilised translational enhancer across a wide range of CF systems, including a 50-fold increase in translational efficiency (Suzuki et al., 2006), and the poly(A) sequence improves mRNA stability, abolishes translational inhibition at higher mRNA concentrations and enhancement of translation (Michel et al., 2000). The combination of these sequences ensures the optimisation of expression and indeed it is seen for the WT CYP3A4 (figure 3.2a). There is a slight increase in WT CYP3A4 yield in CF mixtures supplemented with polymersomes indicating a possible boosting effect through molecular crowding or peptide stabilisation. The lack of expression of Nf14 CYP3A4 however is surprising, given that the WT is readily expressed. With the removal of the membrane spanning N-terminus region to create a more soluble and easier to express CYP3A4 variant indicates that it is this deletion that has possibly led to increased mRNA or peptide instability and degradation.

HeLa and wheat germ CF systems both showed a higher level of overall CYP3A4 expression than the insect CF system (figure 3.2b and c). Interestingly, although both do express the Nf14 CYP3A4, it is expressed to a lower degree than the WT variant indicating the difficulty faced by CF system expressing this construct. The WT constructs being highly expressed, even in the presence of polymersome vesicles (possibly causing a slight increase in expressed protein) is very ideal, showing that CF expression systems can indeed be used for the expression of CYP3A4. The HeLa CF system, being human cell derived, is an ideal system as it is expected to be able to cope with the expression (e.g. ideal codon usage, post-translational modifications and protein translocation into membranes). The pT7CFE1 vector further complements expression with its 5’ UTR from EMCV IRES enhancing translation rates of mRNA (Mikami et al., 2006b), and a 3’ poly(A) sequence.

The wheat germ CF system, like that of the HeLa, can readily express both CYP3A4 constructs although to varying degrees of ease, and is not affected by the presence of polymersomes. Uniquely, this 2-stage (linked transcription and translation), biphasic translation CF system is well suited for expression of CYP3A4. Its specialised vector pEU-E01, contains the E01 translational enhancer sequence upstream of the multiple cloning site, and the biphasic layering of protein translation allows for the gradual exchange of protein expression waste by-products without reaching an inhibitory concentration (Takai and Endo, 2010).

## Summary

CYP3A4 WT and Nf14 constructs were both successfully expressed in multiple CF protein expression systems. Simple and time saving protocols were used to express proteins from DNA in the span of a day, compared to cells based expression which requires multiple days. The yields of CYP3A4 were highest in HeLa and wheat germ CF systems, however, one main advantage of the wheat germ CF over the HeLa system, is the cost of expression. Currently it is still relatively more expensive to use CF systems over standard cell-based expression, and limiting the costs of these systems will help ease limitations on their usage. It is clear that the presence of polymersomes in the CF expression mixture has no major overall effect on protein expression, and possibly even enhancing it slightly. This is one of the major advantages of CF systems, the open-access allowing the control and modification of the expression mixture, grants the ability to introduce biomimetic membranes directly in the presence of protein expression machinery for co-translational insertion and further downstream utilisation. In contrast to the eukaryotic CF expression systems, *E. coli* CF system is incapable of expressing human CYP3A4, while the cell-based expression requires more time-consuming protocols for protein purification and reconstitution.

The next stage is now to express CYP3A4 and the complete monooxygenase system (CPR and C*b5*) in the wheat germ CF systems, chosen for its high-level expression and low cost. Utilising the open access nature of CF systems, the expression mixture will be supplemented with biomimetic vesicles. The ability of the proteins to co-translationally insert into the membranes will be analysed, as well as their propensity for purification and isolation for further downstream applications.

**Chapter 4:**

# Expression and Purification of CYP3A4 Microsomal Monooxygenase Proteopolymersomes

4.1 INTRODUCTION

4.2 AIMS

4.3 EXPERIMENTATION

4.4 CENTRIFUGATION PURIFICATION

4.5 FILTRATION PURIFICATION

4.6 BIOTIN-AVIDIN PURIFICATION

4.7 SEC PURIFICATION

4.8 DISCUSSION

4.9 SUMMARY

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

DLS – Dynamic light scattering

FPLC – Fast protein liquid chromatography

PBS – Phosphate buffered saline

RPM – Revolutions per minute

SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SEM – Standard error of the mean

SPR – Surface plasmon resonance

List of Figures

Figure 4.1: Scheme of proteopolymersome development utilising protein co-translational insertion.

**Figure 4.2:** Percentage of vesicle fluorescence following centrifugation at 10k rpm.

**Figure 4.3:** DLS analysis of polymersome and liposome vesicles sizes at 10K rpm

**Figure 4.4:** Percentage of vesicle fluorescence following centrifugation at 14.5k rpm.

**Figure 4.5:** DLS analysis of polymersome and liposome vesicles sizes with 14.5k rpm

**Figure 4.6:** Fluorescence levels seen in the top (1st fraction) and bottom (2nd fraction) halves of 14.5K rpm centrifugated polymersome sample.

**Figure 4.7:** SYPRO stained SDS-PAGE of vesicle supplemented and centrifugated CF expression mixtures.

**Figure 4.8:** SYPRO stained SDS-PAGE of vesicle supplemented and filtered CF expression mixtures.

**Figure 4.9:** Monitoring of fluorescence levels through the avidin purification of biotinylated polymersomes.

**Figure 4.10:** SYPRO stained SDS-PAGE of avidin purified (proteo)polymersomes.

**Figure 4.11:** SEC purification of (proteo)polymersomes and (proteo)liposomes from crude CF expression product.

## Introduction

Microsomal monooxygenases exist as multi-enzyme systems, and to achieve expression of functional active CYP3A4, simultaneous expression of additional multiple membrane bound proteins is required, namely CPR and C*b5*. To achieve co-translational insertion of the CYP3A4 monooxygenase system (figure 4.1), then CYP3A4 expression is required concurrently with CPR and C*b5* expression, within the same CF expression mixture. With the insertion of CYP3A4, CPR and C*b5* into a single polymersome, it is possible to have protein-protein interactions and a closed system for an electron transport chain to form, a requirement for functional microsomal monooxygenases (Cederbaum, 2015). Typically, CF expression systems aren’t the best suited for over-expression of complex and membrane bound proteins, and the expression of three in a single reaction can add extra stress on the system. The translation of three different proteins requires a broader range of resources and abilities of the CF system, and has a higher risk of incomplete and limited protein translation, peptide mistranslation and aggregation. The capabilities of multi-protein expression in single CF systems have not been fully analysed, currently only one example of multiple protein expression has been published (Cappuccio et al., 2008). In that study, the membrane protein bacteriorhodopsin was co-expressed with a lipoprotein fragment for the solubilisation of proteins onto lipid structures. However, with the microsomal monooxygenase system, the expression of three full length eukaryotic proteins adds more complexity to the system.

The CF system chosen for expression is an important consideration to maximise CYP3A4 functionality. The wheat germ CF system has already been observed with ability to express CYP3A4 efficiently and with high yields, and most importantly in the presence of polymersomes. The use of this system for the co-expression of multiple enzymes has its advantages not offered by the insect and HeLa CF expression systems. Firstly, it’s a linked system allowing separate transcription and translation reactions and therefore while translation must be carried out for all three proteins together, the transcription is carried out individually, lowering the limitation on mRNA production while also increasing the expression yield. The translation stage is biphasic, separating the translation mixture from the “nutrient” mixture allowing the constant exchange of nutrients into the translation mixture coupled with the removal of inhibitory by-products.

|  |
| --- |
|  |
| **Figure 4.1:** Scheme of proteopolymersome development utilising protein co-translational insertion. mRNA transcription occurs in cell-free expression mixture, followed by peptide translation by ribosomes. The membrane binding domain of the peptide inserts into the polymersome membrane as peptide is translated, while the soluble protein domain self-folds in the extra-vesicular environment. |

The first goal for the creation of any proteopolymersome is the embedding of the protein into the membrane. The use of wheat germ CF expression system for the co-translational insertion of membrane proteins has been done previously (Nallani et al., 2011, May et al., 2013, Zapf et al., 2015), and similarly insertion of CYP3A4, CPR and C*b5* into the polymersome membrane has to be confirmed. Standard confirmation methods have utilised antibody binding techniques *e.g.* SPR. An alternative method is the purification of the (proteo)polymersomes and testing for the presence of protein through SDS-PAGE, western blots or flow cytometry. Polymersome purification strategies include filtration; using filtration tubes with membranes of set cut-off values, or affinity; using biotin-avidin binding can isolate the labelled (proteo)polymersome from the impurities of the CF reaction mixture and protein aggregates. One novel tool in the purification of polymersomes is the labelling of membranes using fluorescent markers, *e.g.* lipid-bound rhodamine B, for the tracking of polymersomes throughout the purification process and the development of optimal purification strategies. Such strategies can help elucidate the effectiveness of polymersome purification, highlighting both vesicle loss and yields. The purification of the microsomal monooxygenase proteopolymersome is a vital step in the development of a functional and desirable research tool. Pure preparations allow for the analysis of enzyme activity and provide accurate information on the metabolic activity and function of the enzyme system. Ideally, the best purification method will provide the highest levels of purity proteopolymersomes coupled with the highest yields of protein.

## Aims

The expression of the multi-enzyme CYP3A4 microsomal monooxygenase system in a single CF expression reaction, including the co-expression of the CYP3A4 monooxygenase and CPR only, and the co-expression of CYP3A4, CPR and C*b5* in the presence of polymersome vesicles for the co-translational insertion and creation of microsomal monooxygenase proteopolymersomes. The use of varying techniques, including filtration, centrifugation, surface functionalisation and size exclusion chromatography for the targeted purification of (proteo)polymersomes. Using techniques such as fluorescent labelling of polymersomes and SDS-PAGE, each purification will be analysed for its purity and yields, allowing for the identification of the most favourable purification methods, i.e. high yields and purity of (proteo)polymersomes.

## Experimentation

### Protein co-expression

CYP3A4 was co-expressed with CPR, and C*b5* in the presence of polymersomes, liposomes and also without any vesicles, using the wheat germ CF protein expression system. Transcription for each protein was done separately. The co-translation of the proteins was carried out using varying ratios of mRNA, with 2-fold CPR/C*b5* mRNA to CYP3A4 mRNA. Following translation, protein samples were collected, analysed using SDS-PAGE or Western blot, or used for further purification strategies.

### Fluorescent vesicles

Rhodamine B labelled vesicles were created using the standard film, hydration and extrusion methods, as described in the Methodology section 2.3. These vesicles were dialysed overnight in 1xPBS solution to eliminate any non-associated DPPE-lissamine-rhodamine B. The fluorescence was monitored with excitation at 560 nm and emission at 583 nm.

### Centrifugation purification

Rhodamine B labelled vesicles were monitored for pelleting capacities. The fluorescence of prepared (1 ml) samples of polymersomes and liposomes were measured before, and after centrifugation at 10K and 14K rpm for 30, 60 and 120 minutes. Crude CF expression products containing proteopolymersome and proteoliposomes were also centrifugated, and purity was monitored using SDS-PAGE.

### Filtration purification

Vivaspin filtration columns, with 1000 kDa molecular weight cut-off filters were loaded with crude CF expression products containing proteopolymersomes or proteoliposomes, and centrifugated at 6k rpm for 30 minutes, followed by two PBS buffer (pH 7.4) washes. The retentate was collected and analysed for purity using SDS-PAGE.

### Biotin-avidin purification

Fluorescently labelled, biotinylated polymersomes (figure 2.2) were purified using SoftLink soft release avidin resin. Biotinylation of the vesicles was carried out by mixing biotin conjugated lipid (DSPE-PEG-Biotin) to the polymer prior to the film drying stage. Following the hydration of the dried polymer and the biotin-lipid film, the polymers assemble into vesicles with DSPE-PEG-Biotin incorporated into the membrane. Fluorescence was similarly carried out as biotin labelling using lipid conjugated rhodamine B (DPPE-lissamine-rhodamine B). Fluorescence of the vesicles was monitored throughout all stages of purification for loss and yield of vesicles. Biotinylated proteopolymersomes were also purified using SoftLink soft release avidin resin and analysed for purity on an SDS-PAGE. To maximise efficiency of biotin-avidin purification, altered conditions of the purifications were examined; including the elongation of binding time from 3 hours to overnight, and the use of higher biotin elution concentrations (from 5 mM to 10 mM biotin).

### SEC purification

Crude CF expression product containing fluorescently labelled (proteo)polymersomes were run through a Superdex 200 5/150 GL column on an ÄKTA FPLC system. Separation fractions were collected, the polymersomes elution was tracked using fluorescence measurement, and the protein separation was tracked by Western blot detection.

## Centrifugation purification

Analysis of the centrifugation of fluorescently labelled vesicles has shown the pelleting of PBD-PEO polymersomes and POPC liposomes following centrifugation at 10K and 14K rpm (figure 4.2). Analysis showed that polymersome pelleting levels are around 30% of total vesicles, reaching a plateau after around 2 hours of centrifugation. Liposome vesicles however are considerably more prone to pelleting. They continue to pellet over the 2-hour period, with vesicle fluorescence down nearly 50%. DLS analysis of vesicles, looking at size distribution of the vesicles showed a slight trend to decrease in vesicle size for the polymersomes, and a marked decrease for the liposomes (figure 4.3), indicating a preference for pelleting of larger vesicles as well as shrinkage in size.

Similar analysis of pelleting of vesicles at 14.5K rpm showed similar levels of polymersomes pelleting after 1 and 2 hours, as seen with 10K rpm centrifugation (figure 4.4). This faster centrifugation speed was chosen to try and further separate the polymersomes in the supernatant, and the other contaminants in mixture that can pellet, i.e. lipids, protein aggregates, and CF expression components. When examined at 14.5K rpm, there was a marked increase in percentage of liposome vesicle pelleting, with more than 80% decrease in fluorescence observed.

|  |  |  |  |
| --- | --- | --- | --- |
| **a)** |  | **b)** |  |
| **Figure 4.2:** Percentage of vesicle fluorescence in the supernatant following centrifugation at 10K rpm. Fluorescence analysis of 0.1 ml of polymersomes (10 mg/ml PBD-PEO) and 1 ml of liposomes (1 mg/ml POPC) following centrifugation. **a)** At start, 30 minutes and 60 minutes. **b)** At start, 60 minutes and 120 minutes. Error bars represent the SEM of three independently repeated experiments. | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **a)**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Minutes** | **Size – average**  **(d. nm)** | **Polydispersity**  **index (PdI)** | **Count rate (kcps)** | **Count rate**  **(% of start)** | | **0** | 195.53 | 0.209 | 272.78 | - | | **60** | 179.10 | 0.200 | 204.20 | 74.86 | | **120** | 184.55 | 0.200 | 221.35 | 81.15 | |
| **b)**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Minutes** | **Size – average**  **(d. nm)** | **Polydispersity**  **index (PdI)** | **Count rate**  **(kcps)** | **Count rate**  **(% of start)** | | **0** | 176.05 | 0.121 | 265.55 | - | | **60** | 147.25 | 0.124 | 112.45 | 42.35 | | **120** | 142.00 | 0.113 | 88.90 | 33.48 | |
| **Figure 4.3:** DLS analysis of polymersome and liposome vesicles sizes following centrifugation at 10K rpm. **A)** Polymersome size analysis after 0, 60 and 120 minutes of centrifugation. **B)** Liposome size analysis after 0, 60 and 120 minutes of centrifugation. |

The DLS analysis for the vesicles centrifuged at 14.5K rpm (figure 4.5) showed similar trends as seen with 10K rpm centrifugation, with marked decrease in size distribution and count rate for liposomes, while only minor decreases for polymersome vesicles.

It is notable that with 14.5K rpm centrifugation, polymersomes do begin to pellet. This is evident through the higher levels of fluorescence in the lower half of tube compared to the top half. Polymersomes (1 ml volume), when centrifuged at 14.5K rpm, showed almost half the levels of fluorescence at the top half (500 µl), while the bottom half (500 µl) had similar fluorescence levels to that seen at the start (figure 4.6).

|  |
| --- |
|  |
| **Figure 4.4:** Percentage of vesicle fluorescence in the supernatant following centrifugation at 14.5K rpm. Fluorescence analysis of 0.1 ml of polymersomes (10 mg/ml PBD-PEO) and 1 ml of liposomes (1 mg/ml POPC) following centrifugation. Error bars represent the SEM of three independently repeated experiments |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **a)**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Minutes** | **Size – average**  **(d. nm)** | **Polydispersity index (PdI)** | **Count rate**  **(kcps)** | **Count rate**  **(% of start)** | | **0** | 159.8 | 0.149 | 679.3 | - | | **60** | 146.4 | 0.141 | 537.2 | 79.1 | | **120** | 144.3 | 0.174 | 557.7 | 82.1 | |  |  |  |  |  | |
| **b)**   |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Minutes** | **Size - average**  **(d. nm)** | **Polydispersity index (PdI)** | **Count rate**  **(kcps)** | **Count rate**  **(% of start)** | | **0** | 176.1 | 0.121 | 265.6 | - | | **60** | 151.1 | 0.071 | 45.2 | 17.0 | | **120** | 147.5 | 0.080 | 49.6 | 18.7 | |
| **Figure 4.5:** DLS analysis of polymersome and liposome vesicles sizes following centrifugation at 14.5K rpm. **A)** polymersome size distribution after 0, 60 and 120 minutes of centrifugation at 14.5K rpm. **B)** Liposome size distribution after 0, 60 and 120 minutes of centrifugation at 14.5K rpm. |

|  |
| --- |
|  |
| **Figure 4.6:** Fluorescence levels seen in the top (1st fraction) and bottom (2nd fraction) halves of the supernatant following 14.5K rpm centrifugated polymersome sample. Fluorescence analysis of 0.1 ml of polymersomes (10 mg/ml PBD-PEO) following centrifugation. Error bars represent the SEM of three independently repeated experiments. |

Subsequently, purification of 3-enzyme (CYP3A4, CPR and C*b5*) (proteo)polymersome and (proteo)liposome vesicles from the CF expression mixture was carried out using centrifugation of the crude CF expression product. SDS-PAGE showed that most of protein impurities present in the crude expression extract do remain present in the supernatant post centrifugation at 10K rpm for 1 hour (figure 4.7). The pellet fraction showed very little presence of any proteins.

|  |
| --- |
|  |
| **Figure 4.7:** SYPRO stained SDS-PAGE of vesicle supplemented and centrifugated CF expression mixtures (3-enzyme CYP3A4, CPR and C*b5* monooxygenase system). Polymersome supplemented CF expression mixtures (lanes 1-3), and liposome supplemented expression mixture (lanes 4-6), showing crude expression product (lanes 1 & 4), supernatant (lanes 2 & 5), and pellet fraction (lanes 3 & 6).  CPR ~78 kDa, CYP3A4 ~59 kDa, C*b5* ~17 kDa. |

## Filtration purification

Vivaspin columns with a 1000 kDa molecular weight cut-off filter were trialed for the filtration of (proteo)polymersomes from the crude CF expression product, and the purity analysed using SDS-PAGE. Vivaspin filtration of the CF expression mixtures showed the loss of a large amount of the protein impurities present (figure 4.8), and higher purity levels than observed through centrifugation alone. However, large amount of impurities was still observed.

|  |
| --- |
|  |
| **Figure 4.8:** SYPRO stained SDS-PAGE of vesicle supplemented and filtered CF expression mixtures (3-enzyme CYP3A4, CPR and C*b5* monooxygenase system). Polymersome supplemented CF expression mixtures (lanes 1-2), and liposome supplemented expression mixture (lanes 3-4), showing crude expression mixture (lanes 1 & 3) and vivaspin retentate (lanes 2 & 4).  CPR ~78 kDa, CYP3A4 ~59 kDa, C*b5* ~17 kDa. |

## Biotin-avidin purification

Fluorescently labelled vesicles were used to monitor avidin purification of membrane biotinylated vesicles (figure 2.2). The purification was split into three parts, the binding onto the avidin resin, the washing with PBS solution and the elution off the resin with biotin. Fluorescence was monitored at all stages of fluorescence indicating the successful purification (figure 4.9). The fluorescence levels showed that over a third of vesicles present at start of purification did not bind to the avidin resin and were washed away. The first elution with avidin showed higher levels of fluorescence than at start indicating increase in polymersome concentration (figure 4.9a). However, when accounting for the total volume of sample, the total amount of fluorescence in the first elute is about 40% of the start fluorescence, indicating a 60% loss of polymersome (lost in the unbound fraction or uneluted and bound to the avidin resin). Purification of biotinylated liposomes was attempted, but was unsuccessful.

|  |  |  |  |
| --- | --- | --- | --- |
| **a)** |  | **b)** |  |
| **Figure 4.9:** Monitoring of fluorescence levels through the avidin purification of biotinylated fluorescent polymersomes. The fluorescence is monitored at the start (total), the unbound to the avidin resin (unbound), and those eluted off the avidin resin (E1 and E2). **a)** Fluorescence levels in the collected samples (fluorescence/0.1 ml). **b)** Polymersome fluorescence levels taking into account sample volume (total fluorescence). | | | |

Avidin purification of biotinylated (proteo)polymersomes from CF expression mixture was carried out. Samples were run on a SDS-PAGE for analysis of purity (figure 4.10). Purity of the avidin purified (proteo)polymersome sample is much higher than that observed with either centrifugation or filtration purification strategies, with the microsomal monooxygenase proteins; CYP3A4, CPR and C*b5* are clearly visible.

|  |
| --- |
|  |
| **Figure 4.1**0**:** SYPRO stained SDS-PAGE of avidin purified (proteo)polymersomes. Total CF expression mixture (3-enzyme system, lane 1), 2-enzyme (proteo)polymersome avidin purification (lane 2), and 3-enzyme (proteo)polymersome avidin purification (lane 3).  CPR ~78 kDa, CYP3A4 ~59 kDa, C*b5* ~17 kDa. |

## SEC purification

The resolution of separation by the column was too diffuse, and the crude CF expression products containing fluorescently labelled (proteo)polymersomes and (proteo)liposomes were separated across a large volume. The tracking of vesicle fluorescence indicated the separation of vesicles across a volume of 0.5 ml (figure 4.11a), and separation of protein across 2.5 ml, a 25-fold dilution (figure 4.11b).

|  |
| --- |
| **a)** |
| **b)**   |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Separation volume (ml)** | 1.15 | 1.35 | 1.55 | 1.75 | 1.95 | 2.15 | 2.35 | 2.55 | 2.75 | |
| **Figure 4.11:** SEC purification of (proteo)polymersomes (blue) and (proteo)liposomes (orange) from crude CF expression product. **a)** Separation of (proteo)vesicles shown by fluorescence tracking (0.1 ml fraction collection). **b)** Western blot of separated microsomal monooxygenase proteins from polymersome supplemented CF expression mixture.  CPR ~78 kDa, CYP3A4 ~59 kDa |

## Discussion

When working with microsomal monooxygenases preparations, as with any enzyme preparations, the purity is important. Expressed proteins, whether from cell-based expression systems or using CF expression systems, will always require a purification processes. The use of CF expression systems supplemented with polymersome (or liposomes) vesicles requires a purification method to isolate the expressed proteins, especially those embedded on the polymersome membrane, from the surrounding impurities such as the native CF expression proteins and the other chemicals present. With CF expression systems, a major advantage in purification is the openness of the system. Easy access to (proteo)polymersomes allow for the development and utilisation of simple and quick purification strategies. Two main strategies for the purification of proteopolymersomes exist, either the purification of proteins or the purification of the vesicles. The strategy focused on here is the targeted purification of the (proteo)polymersome vesicles. This avoids the possibility of purifying undesired target proteins, such as those in aggregates or those not anchored on the polymersome.

The simple centrifugation purification strategy was first utilised for (proteo)polymersome separation. Typically, centrifugation has been used for the pelleting of expressed protein aggregates and membrane fractions. Analysis of vesicle pelleting indicated as expected the liposome pelleting (figures 4.2 and 4.4), and indeed it has been widely utilised as a purification strategy for proteoliposomes (Liguori and Lenormand, 2009, Liguori et al., 2010, Moritani et al., 2010). Conversely however polymersomes are more resistant to pelleting, showing only small amounts of vesicle loss to the pellet (figure 4.2). DLS analysis indicated polymersome and liposome vesicle sizes profile as expected in the 100 – 200 nm range, however the liposome showed size shrinkage following centrifugation. The vesicle count rate for polymersomes and liposomes follows a very similar trend to the fluorescence readings following centrifugation, and confirms the pelleting of the vesicles. The combination of fluorescence readings and DLS analysis confirms the propensity of liposomes for pelleting during centrifugation causing decrease in vesicle count in the supernatant as well as shrinkage in size, while polymersomes are more opposed to pelleting.

This difference in vesicle pelleting characteristics was taken advantage of to purify (proteo)polymersomes from its CF expression mixture protein impurities and the unanchored microsomal enzyme aggregates. Ideally, with centrifugation, the supernatant fraction would be pure enough to visualise the expressed protein bands while the impurities and aggregates would only be visible in the pellet fraction. However, centrifugation was not sufficient for (proteo)polymersome purification (figure 4.7), with the majority of the protein impurities present in the crude CF expression product still remaining in the supernatant fraction following centrifugation for 1 hour. For (proteo)liposomes, centrifugation was a more effective method of purification, as the pellet fraction, where the majority of liposomes are expected contains much less impurities. However, there is still small amount of CF expression mixture proteins that do pellet and contaminate this fraction. Further use of centrifugation as a method of purification of (proteo)polymersomes will require modifications of the strategies used including use of higher speeds and longer durations, however it seems this strategy is limited in its capabilities for (proteo)polymersome purification as it not directly target the vesicles, rather utilise the physical characteristics of the polymersomes and its surrounding environment.

The use of filtration membranes, in this case Vivaspin columns, also aims to take advantage of the physical properties of the polymersomes and liposomes. DLS analysis of polymersomes and liposomes showed vesicle sizes in the range of 100-200 nm, larger than most proteins. The aim was to utilise the (1000 kDa molecular weight) filter membrane for the retention of the (proteo)polymersomes and (proteo)liposomes, while allowing the flow-through of unbound aggregates and other impurities. While in general the retention of vesicles was possible, there was still a sizeable amount of impurities retained despite numerous repeated washes with PBS buffer (figure 4.8).

Avidin purification of the biotin labelled vesicles was a much more successful strategy for the purification of (proteo)polymersomes. The targeted labelling and purification of polymersomes is a much more specific purification strategy than centrifugation or filtration. The resulting preparation is very pure, seen in the SDS-PAGE (figure 4.10), where only the expression protein bands are visible. This successful avidin purification for polymersomes is also clear indication of the co-translational insertion of protein into the polymersome membrane. Since avidin only targets biotinylated polymersomes, only proteins anchored on polymersomes, from co-translational insertion, are present. The avidin purification of biotin labelled (proteo)liposomes was attempted, however no protein presence was observed. Limiting factors for the avidin purification strategy however do still exist. With roughly only 40% of vesicles recovered, albeit at a higher concentration than at the start, it is a low yield. Other avenues for avidin purification can be examined for optimisation of yields. The biotin-avidin bond is one of the strongest non-covalent bonds (Kd = 10-15 M), and the bond is uninterrupted by extremes in pH, temperature or denaturing agents, requiring harsh conditions for dissociation that would denature proteins. Avidin-biotin purifications are therefore complicated and not ideal, requiring high biotin concentrations and multiple elutions. Currently the avidin purification protocol used has a loss of around 20% of vesicles that remain attached to the avidin resin after first elution. The use of a biotin alternative, such as desthiobiotin to bind to avidin with similar specificity, however the bond is much more reversible and advantageous.

The use of SEC separation methods while viable in theory, are not an ideal purification method. The polymersome separation profile is too diffuse, and similarly the proteins were separated across a wide profile. With a 100 μl sample run through the SEC column separated across 1.5 ml volume. This is highly disadvantageous for CF expressed proteins since the low expression already obtained with the CF system. Possible optimisation of SEC separation can be achieved through use of columns with higher separation resolution, leading to the narrowing of the polymersome peak and its separation into a smaller volume.

## Summary

The multiple attempted purification strategies have their own degrees of success. Centrifugation seemingly lead to indistinguishable improvements of (proteo)polymersome purity, while similarly the filtration of the CF expression mixture only fared slightly better. The SEC purification is not currently viable as is and requires significant changes to improve the resolution of purification. Avidin purification of biotinylated vesicles is currently the best purification strategy examined, offering excellent purity of the microsomal monooxygenase enzymes present on the polymersome membrane, with complete removal of the CF expression system impurities and loss of all undesired expression aggregates. Additionally, and quite importantly, avidin purification lead to increased concentration of (proteo)polymersomes, albeit in a smaller volume, conserving the low yields usually obtained with CF expressed systems. There are limitations to the purification analysis carried out in this chapter, namely the state of the actual purification product.

The analysis carried out give no indication towards the folding state proteins, and their functionality. All purification protocols exert their own stresses on enzymatic systems which can lead to denaturing of proteins and their misfolding causing loss of function. The true test of a successful purification is through functionality, and the various purification techniques tested here will have to be tested for functionality, to assess the effect of purification on the levels of activity if any, and to ensure no loss of functionality. The avidin purification in particular is of special interest as it allows quantification of the specific activity of the CYP3A4 enzyme and the elucidation of how active and pure the elutant product is.

**Chapter 5:**

# Testing the Activity of the CYP3A4 Proteopolymersome

5.1 INTRODUCTION

5.2 AIMS

5.3 EXPERIMENTATION

5.4 TOTAL CF CYP3A4 PROTEOPOLYMERSOME

5.5 ALTERING CONDITIONS OF EXPRESSION

5.6 CENTRIFUGED CF CYP3A4 MONOOXYGENASE ACTIVITY

5.7 FILTERED CF CYP3A4 MONOOXYGENASE ACTIVITY

5.8 AVIDIN PURIFIED CYP3A4 PROTEOPOLYMERSOME ACTIVITY

5.9 DISCUSSION

5.10 SUMMARY

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

FAD – Flavin adenine dinucleotide

FMN – Flavin mononucleotide

PTM – Post-translational modification

SEM – Standard error of the mean

List of Tables

**Table 5.1:** The mass and mass ratio of proteins present in the CYP3A4 monooxygenase proteopolymersome and the Baculosome.

List of Figures

**Figure 5.1:** Western blot of total 3-enzyme CYP3A4 microsomal monooxygenase system expression in wheat germ CF protein expression system.

**Figure 5.2:** Activity levels of the 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system.

**Figure 5.3:** Activity levels of the 2-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system.

**Figure 5.4:** Activity of wheat germ CF expressed CYP3A4 microsomal monooxygenase system under different expression conditions.

**Figure 5.5:** Western blot of centrifuged 3-enzyme CYP3A4 microsomal monooxygenase system expression in wheat germ CF protein expression system.

**Figure 5.6:** Activity levels of the centrifuged 2-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system.

**Figure 5.7:** Activity levels of the centrifuged 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system.

**Figure 5.8:** Activity levels of the filtered 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system.

**Figure 5.9:** SYPRO stained SDS-PAGE of avidin purified CYP3A4 monooxygenase proteopolymersome.

**Figure 5.10:** Activity levels of the avidin purified 2- and 3-enzyme CYP3A4 microsomal monooxygenase proteopolymersome.

**Figure 5.11:** Analysis of the purity and activity of Baculosome preparation.

## Introduction

Functionality, coupled with purity, are two of the most important parameters in the development of enzyme preparations. To achieve functionality in microsomal monooxygenase proteopolymersomes, there needs to be co-operation in this multi-enzyme system to form a linked electron transport chain from CPR and C*b5* to CYP3A4. It is important to ensure that the enzymes present on the membrane are capable of folding in a conformation that allows close interactions between each other. Ideally this will be most likely to happen with enzymes anchored on the polymersome membrane, while the unbound proteins and aggregates will present with no or lower levels of activity.

A lot of factors within the CF expression mixtures and the enzymes themselves can affect the functionality of the produced microsomal monooxygenase. The presence of polymersomes (or liposomes) in the expression mixture should provide the stabilising environment for the enzymes, and as such without their supplementation into the CF mixture, less activity and more inactive protein aggregates should be observed. The enzyme conditions themselves could be altered for optimisation of activity levels, including optimisation of CPR and C*b5* expression levels compared to CYP3A4. CYP3A4, as well as CPR and C*b5*, require cofactor supplementation without which no activity should be observed. The localisation of the enzymes on the proteopolymersome can be examined by comparing activity of the CYP3A4 WT to the Nf14 mutant. This mutant lacking the trans-membrane domain cannot anchor itself to the polymersome membrane, however it has been shown to still interact and bind to lipid membranes. Any activity observed with it is an indication of its possible ability to bind to polymer-based membranes.

Attempting separate expression of each component of the CYP3A4 microsomal monooxygenase system is an avenue for the examination of the interactions of these enzymes. Separately produced proteopolymersomes should have only one individual member of the monooxygenase system on vesicle membranes, and the subsequent mixing of such samples will require close interactions of the proteopolymersomes to allow for electron transport.

For the CYP3A4 microsomal monooxygenase proteopolymersome, sample preparations have been created, from the very impure proteopolymersomes in crude CF expression product, to the slightly purer centrifuged and filtered proteopolymersomes in CF expression mixtures, and the highly pure avidin purified proteopolymersome preparations. The activity of all these different preparations needs to be examined to monitor and understand the effect of purification, if any, on functionality. The activity of the CYP3A4 microsomal proteopolymersomes in the crude CF expression product will be used as a standard baseline for activity, and from that the activity of purified preparations can be compared to.

Following establishing the functionality of the CYP3A4 microsomal monooxygenase proteopolymersome, further analysis can be carried out including measuring the specific activity of the CYP3A4 enzyme. The specific activity is reliant on the enzymes ability to generate the product per mg of protein and therefore it can be used as a marker of purity for correctly anchored, folded, and active enzyme.

## Aims

To test the capability of wheat germ CF protein expression system to express multiple functional enzymes including microsomal monooxygenase enzymes, and to test that activity with and without polymersome supplementation. To compare the activity levels observed with polymersome and liposome supplemented CF expression mixtures. To identify the importance of CPR and C*b5* enzymes in achieving an active CYP3A4 microsomal monooxygenase system, as well as the role of cofactor supplementation during the CF expression stage on the activity of all three enzymes. Using separate expression, as well the expression of trans-membrane domain deficient CYP3A4 to examine the localisation of the expressed enzymes, and to assess the tightness of interaction between them. The effect of proteopolymersome purification on microsomal monooxygenase functionality will be monitored for any decrease in activity, as well as assessing the specific activity of CYP3A4 in the pure microsomal monooxygenase preparations, with the aim of obtaining very pure and highly active proteopolymersome preparations.

## Experimentation

### Protein co-expression

CYP3A4 was co-expressed with CPR for 2-enzyme microsomal monooxygenase system, and with CPR and C*b5* for 3-enzyme microsomal monooxygenase system using the wheat germ CF protein expression system. Polymersome and liposome supplementation was carried out as described in the Methodology, section 2.4. Transcription for each protein was done separately. The co-translation of the proteins was carried out using varying ratios of mRNA, with 2-fold CPR/C*b5* mRNA to CYP3A4 mRNA.

### Activity analysis

Vivid blue CYP3A4 screening kit was used to test CYP3A4 monooxygenase functionality, as described in the Methodology, section 2.7. Fluorescence of the BOMCC product (ex: 415, em:460) was monitored over 2 hours. Experiments were done in triplicates and the SEM was calculated and plotted as error bars.

## Total CF CYP3A4 monooxygenase activity

The 2- and 3-enzyme CYP3A4 microsomal monooxygenase system were successfully expressed using wheat germ CF expression system, supplemented with polymersomes and liposomes (figure 5.1).

Both the 2- and 3-enzyme CYP3A4 microsomal monooxygenase system were shown to be functional (figure 5.2 and 5.3 respectively). Polymersome supplemented CF expression systems showed highest levels of activity, performing slightly better than liposome supplemented CF expression systems. CYP3A4 activity was lower without any vesicle supplementation. The 2-enzyme CYP3A4 microsomal monooxygenase system showed significantly higher levels of activity compared to the 3-enzyme system, at almost five times higher.

|  |
| --- |
|  |
| **Figure 5.1:** Western blot of total 3-enzyme CYP3A4 microsomal monooxygenase system expression in wheat germ CF protein expression system. CYP3A4, CPR and C*b5* were successfully expressed using wheat germ CF protein expression system. 3-enzyme expression in polymersome supplemented protein expression mixture (lane 1). 2-enzyme expression, without CYP3A4, in polymersome supplemented protein expression mixture (lane 2). 3-enzyme expression in liposome supplemented protein expression mixture (lane 3). 3-enzyme expression in non-vesicle supplemented protein expression mixture (lane 4).  CPR ~78 kDa, CYP3A4 ~59 kDa, C*b5* ~17 kDa. |

|  |
| --- |
|  |
| **Figure 5.2:** Activity levels of the 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system. Blue: polymersome supplemented CF expression system, Orange: liposome supplemented CF expression system, Grey: non-vesicle supplemented CF expression system, and yellow: no CYP3A4 expression in polymersome supplemented CF protein expression system. Error bars represent the SEM of three independently repeated experiments. |

|  |
| --- |
|  |
| **Figure 5.3:** Activity levels of the 2-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system. Blue: polymersome supplemented CF expression system, Orange: liposome supplemented CF expression system, Grey: non-vesicle supplemented CF expression system, and yellow: no CYP3A4 expression in polymersome supplemented CF protein expression system. Error bars represent the S.E.M. of three independently repeated experiments. |
|  |

## Altering conditions of expression

Alteration of the conditions of expression had a marked effect on the functionality of the CYP3A4 microsomal monooxygenase system (figure 5.4). As seen in figures 5.2 and 5.3, the 3-enzyme CYP3A4 monooxygenase system had less than a quarter of the activity levels of the 2-enzyme system. Another 2-enzyme system, composed of CYP3A4 and C*b5* only, showed no levels of activity (figure 5.4, bar 3). Similarly, no significant activity was observed for the 2-enzyme CYP3A4 monooxygenase system expressed without the hemin, FMN and FAD cofactor supplementations (figure 5.4, bar 4). Minimal levels of activity were observed when WT CYP3A4 was substituted for the mutant, trans-membrane domain deficient CYP3A4 Nf14 (figure 5.4, bar 5. Similarly, when CYP3A4, CPR and C*b5* enzymes were independently expressed and subsequently mixed, low levels of activity were observed (figure 5.4, bar 6). Caution must be taken however as while the first five samples (figure 5.4, bars 1-5) were prepared with only 50 µl of mRNA, sample six (figure 5.4, bar 6) was prepared with 150 µl of mRNA.

## Centrifuged CF CYP3A4 monooxygenase activity

The Western blot of the centrifuged CYP3A4 monooxygenase system showed that the majority of the expressed proteins remain in the supernatant fraction, regardless of vesicle presence, with barely detectable levels of proteins present in the pellet fraction (figure 5.5). The activity of the samples however paints a different picture. The activity levels for the pellet fraction is higher than in the supernatant, for both the 2- and 3-enzyme CYP3A4 monooxygenase system (figure 5.6 and 5.7 respectively). While the activity levels in the supernatant fraction shows a significant decrease from the crude total CF CYP3A4 monooxygenase samples of around 40 – 50% of the activity, the pellet fraction maintains roughly around the same levels of activity.

|  |
| --- |
| **1 2 3 4 5 6** |
| **Figure 5.4:** Activity of wheat germ CF expressed CYP3A4 microsomal monooxygenase system under different expression conditions (activity/0.05 ml of prepared sample). Bar 1: 2-enzyme CYP3A4 monooxygenase system (. bar 2: 3-enzyme CYP3A4 monooxygenase system, bar 3: CPR deficient CYP3A4 monooxygenase system, bar 4: 2-enzyme CYP3A4 monooxygenase system no cofactor (heme, FAD, FMN) supplementation, bar 5: 2-enzyme CYP3A4 Nf14 monooxygenase system. Bar 6: Independently expressed and mixed 3-enzyme CYP3A4 monooxygenase system. Error bars represent the SEM of three independently repeated experiments. |

|  |
| --- |
|  |
| **Figure 5.5:** Western blot of centrifuged 3-enzyme CYP3A4 microsomal monooxygenase system expression in wheat germ CF protein expression system. The Western blot confirms that the majority of expressed the microsomal monooxygenase proteins are present in the supernatant fraction. Supernatant (lane 1) and pellet (lane 2) fractions of 3-enzyme expression in polymersome supplemented protein expression mixture. Supernatant (lane 3) and pellet (lane 4) fractions of 2-enzyme expression, without CYP3A4, in polymersome supplemented protein expression mixture. Supernatant (lane 5) and pellet (lane 6) fractions of 3-enzyme expression in liposome supplemented protein expression mixture. Supernatant (lane 7) and pellet (lane 8) fractions of 3-enzyme expression in non-vesicle supplemented protein expression mixture.  CPR ~78 kDa, CYP3A4 ~59 kDa, C*b5* ~17 kDa. |

|  |
| --- |
| **a)** |
|  |
| **b)** |
| **Figure 5.6:** Activity levels of the centrifuged 2-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system. **a)** Activity levels in the supernatant fraction. **b)** Activity levels in the pellet fraction. Blue: polymersome supplemented CF expression system, Orange: liposome supplemented CF expression system, Grey: non-vesicle supplemented CF expression system, and yellow: no CYP3A4 expression in polymersome supplemented CF protein expression system. Error bars represent the SEM of three independently repeated experiments. |

|  |
| --- |
| **a)** |
|  |
| **b)** |
| **Figure 5.7:** Activity levels of the centrifuged 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system. **a)** Activity levels in the supernatant fraction. **b)** Activity levels in the pellet fraction. Blue: polymersome supplemented CF expression system, Orange: liposome supplemented CF expression system, Grey: non-vesicle supplemented CF expression system, and yellow: no CYP3A4 expression in polymersome supplemented CF protein expression system. Error bars represent the SEM of three independently repeated experiments. |

There are variations in the activity levels of the different samples between the supernatant and the pellet fractions. While in the supernatant fractions, generally the polymersome supplemented CYP3A4 monooxygenase systems have the highest levels of activity (figure 5.6a and 5.7a), in the pellet fractions, the liposome-supplemented and non-vesicle supplemented samples have high levels of activity (figure 5.7b). This is an outcome of the varying pelleting tendencies of the different vesicles and the unbound aggregates. It is worth noting that the total values of the activity seen in the supernatant fraction and the pellet fractions (four-fold concentrated) in figures 5.6 and 5.7, correspond closely to the total activity values observed in figures 5.3 and 5.2, respectively.

## Filtered CF CYP3A4 monooxygenase activity

Testing of the activity levels of the vivaspin filtered CYP3A4 monooxygenase samples was unsuccessful. Activity levels form the polymersome and liposome supplemented CYP3A4 monooxygenase CF expression mixtures were as low as the CYP3A4 negative samples (figure 5.8).

|  |
| --- |
|  |
| **Figure 5.8:** Activity levels of the filtered 3-enzyme CYP3A4 microsomal monooxygenase system expressed in wheat germ CF protein expression system. Blue: polymersome supplemented CF expression system, Orange: liposome supplemented CF expression system, and Grey: non-vesicle supplemented CF expression system. Error bars represent the SEM of three independently repeated experiments. |

## Avidin purified CYP3A4 proteopolymersome activity

Avidin purification of the 2- and 3-enzyme CYP3A4 monooxygenase system was successful, with high levels of purity shown on SDS-PAGE (figure 5.9). Densitometry analysis of the gels, using the BSA standards, showed excess of CPR and C*b5* to CYP3A4 (table 5.1). In the 3-enzyme CYP3A4 monooxygenase system expression, twice as much mRNA of CPR and C*b5* as was used for CYP3A4, and the resulting protein expression followed similar trend of 2.3 µg and 1.9 µg of CPR and C*b5* proteins respectively per µg of expressed CYP3A4. However, with the 2-enzyme monooxygenase system, while the ratios of mRNA used remains the same, the ratios of proteins expressed changed, almost doubling up to 3.8 µg of CPR expressed per µg of CYP3A4 expressed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 5.1:** The mass and mass ratio of proteins present in the CYP3A4 monooxygenase proteopolymersome and in the CYP3A4 Baculosome. | | | | | | |
| **Protein** | **2-enzyme proteopolymersome** | | **3-enzyme proteopolymersome** | | **Baculosomes**  **(20-fold dilution)** | |
|  | **µg/ml** | **Mass ratio** | **µg/ml** | **Mass ratio** | **µg/ml** | **Mass ratio** |
| **CPR** | 39.8 | 3.8 | 12.9 | 2.3 | 11.2 | 2.11 |
| **CYP3A4** | 7.8 | 1 | 5.7 | 1 | 5.3 | 1 |
| **C*b5*** | – | – | 11.2 | 1.9 | 3.9 | 0.75 |

Similar activity trend, as seen in the total CYP3A4 monooxygenase system, is seen with the avidin purified proteopolymersomes. the 2-enzyme proteopolymersome has a much higher level of activity than that of the 3-enzmye (figure 5.10).

|  |
| --- |
| **a)**  **CPR**  **CYP3A4**  **25 15 5**  **1 2** |
| **b)**  **CPR**  **CYP3A4**  **C*b5***  **25 15 5**  **1 2** |
| **Figure 5.9:** SYPRO stained SDS-PAGE of avidin purified CYP3A4 monooxygenase proteopolymersome. High levels of expressed protein purity is observed for both proteopolymersome preparations. **a)** Avidin purified 2-enzyme CYP3A4 monooxygenase proteopolymersome (lane 1), and with no CYP3A4 expression (lane 2). With BSA standard (25, 15 and 5 µg/ml). **b)** Avidin purified 3-enzyme CYP3A4 monooxygenase proteopolymersome (lane 1), and with no CYP3A4 expression (lane 2). With BSA standard (25, 15 and 5 µg/ml).  CPR ~78 kDa, CYP3A4 ~59 kDa, Cb5 ~17 kDa. |

|  |
| --- |
|  |
| **Figure 5.10:** Activity levels of the avidin purified 2- and 3-enzyme CYP3A4 microsomal monooxygenase proteopolymersome. Blue: 2-enzyme CYP3A4 monooxygenase proteopolymersome, Orange: 3-enzyme CYP3A4 monooxygenase proteopolymersome, Grey: CYP3A4 deficient proteopolymersome. Error bars represent the SEM of three independently repeated experiments. |

The high levels of purity, as shown in figure 5.9, allow for the calculation of the specific activity of CYP3A4, using the following equation:

Where *T* is reaction time (Tend – Tstart), and *P* is the amount of protein in the sample. A fluorescent standard curve was used to calculate the BOMCC product formation concentration (appendix 4). The specific CYP3A4 enzyme activity was calculated to be 5.4 U/mg (µmol min-1 mg-1) for the BOMCC substrate at 25oC, pH 7.2, while the total enzyme activity was calculated to be 0.02 U/ml (µmol min-1 ml-1) for the BOMCC substrate at 25oC, pH 7.2. The Baculosomes, a commercial CYP3A4 monooxygenase microsomal preparation (from Baculovirus infected insect cells), was also compared to the proteopolymersome sample. Baculosome exhibited similar levels of purity yet much higher levels of activity (figure 5.11).

|  |
| --- |
| **a)** |
|  |
| **b)** |
| **Figure 5.11:** Analysis of the purity and activity of Baculosome preparation. **a)** SYPRO stained SDS-PAGE of Baculosomes, with BSA standard (25, 15 and 5 µg/ml). **b)** Measured activity levels of 10 µl of 20-fold dilution CYP3A4 Baculosomes. Error bars represent the SEM of three independently repeated experiments. |

Analysis of the specific activity of the CYP3A4 enzyme in the commercial CYP3A4 Baculosome preparation was much higher, at 680 U/mg.

## Discussion

Functionality analysis of the wheat germ CF expressed CYP3A4 microsomal monooxygenase system has shown that it is indeed active, and the presence of vesicles, both polymersomes and liposomes, does increase the activity levels of the system. surprisingly however there is a high level of activity present in the non-vesicle supplemented sample, at around 75% of the vesicle supplemented samples, indicating possible rearrangement of the enzyme on native wheat germ CF containing membranes or possibly as active aggregates. The increase of activity with vesicle supplementation however is as expected with the increased availability of stabilising membrane environments, there is increased number of correctly inserted and folded enzymes, and therefore improved functionality. However, unlike most published data, it seems the additional expression of C*b5* leads to a decrease in CYP3A4 monooxygenase activity. This is most likely due to the limitations of the wheat germ CF system rather than the monooxygenase system itself. The over-expression of multiple complex proteins can stress the system and lead to incomplete and misfolded protein expression.

Centrifugation of the expressed CYP3A4 monooxygenase system, although did not lead to significant increase in purity, however it did cause a significant decrease in activity levels, reaching almost a 60% loss of functionality (figure 5.6 and 5.7). The pellet fraction on the other hand seemingly had very small amount of protein present however the activity levels was almost the same as that seen with total CYP3A4 monooxygenase system. These results hint towards limited pelleting of any proteins, whether on the polymersomes or on other native membranes or as aggregates, other than a small number of functional enzymes on non-polymersome membranes that have sufficiently high levels of activity.

The total loss of all activity following Vivaspin purification was unexpected. Although this filtration method did improve the purity of the sample more than centrifugation. From the Western blot, it is clear to see that the enzymes do remain on the polymersome in the filtration retentate, and that all aggregated proteins in the absence of polymersomes pass through. This indicates the loss of activity is due to the properties of the enzymes, *e.g.* protein unfolding or loss of cofactors during the filtration and wash stages, rather than the loss of proteopolymersomes or loss of the polymersome structural integrity.

The avidin purified CYP3A4 monooxygenase proteopolymersome is the most promising CF produced preparation. The purity is excellent and clear of any protein impurities. while the activity of these purified proteopolymersomes can reach high levels, the specific activity has more room for improvement. When compared to the commercial Baculosome (microsomes prepared from Baculovirus infected insect cells) preparation, there is room for improvement. The activity pattern seen with the proteopolymersomes are like those seen previously in the non-purified samples, with the 2-enzyme proteopolymersome system giving nearly eight times higher levels of activity. The specific activity of the CYP3A4 proteopolymersome improves significantly simply by expressing CPR only with CYP3A4, without C*b5*. This significant difference in both the total activity and specific activity can be attributed to several factors. The main factor affecting the activity of the CYP3A4 monooxygenase system is its multi-component composition. It is known that microsomal monooxygenase systems are heteromeric multi-enzyme systems with CPR and C*b5* present in excess. Having excess of reductases has shown to increase electron transport to CYP3A4 and increase activity. Work has been carried out to identify the ideal ratio of CYP3A4 to CPR and C*b5* (Lee and Goldstein, 2012), and it has been shown that can be up to 8 CPR and 16 C*b5* per CYP3A4. In comparison, the proteopolymersomes were prepared with only twice the concentration of CPR mRNA to CYP3A4 mRNA. The improved specific and total activity seen in the 2-enzyme CYP3A4 enzyme system could be due to the doubling of CPR expression levels, from 2-fold the level of CYP3A4 expression seen with the 3-enzyme CYP3A4 system, to nearly 4-fold more expression. Optimising CPR availability to boost the specific activity of CYP3A4 proteopolymersome, as well as to co-express with C*b5* in excess is a challenge for CF expression systems, but tackling it is required to be fully able to compete with commercial cell-based CYP3A4 preparations such as the Baculosomes.

Since with the avidin purified proteopolymersomes all the purified proteins are associated with polymersomes in some capacity, it is generally acceptable to assume that they have indeed successfully anchored into the membrane. However, it is no indication of the folding state of the proteins, and it is also possible that a small amount of the proteins have associated with the polymersomes incorrectly or partially and therefore lowering the specific activity. The wheat germ CF expression system is known to have limited protein processing and PTM capabilities, and as such the improvement of protein folding can be limited by the chosen expression system and alternative CF expression systems might have to be examined for significant improvement of protein folding and incorporation on the polymersome membrane. More complex and capable CF systems such as insect and HeLa are known for providing better protein folding and processing which can aid in developing functional proteopolymersomes.

Alternative strategies can be developed for the improvement of the CYP3A4 proteopolymersomes activity. Reconstitution methods can be used to fine-tune and control the amounts of enzymes present on polymersome membranes. The use of CF systems, supplemented with detergents, *e.g.* Fos-choline 14, can be used for the stable expression of purifiable protein. These proteins can be purified and reconstituted on polymersomes using standard reconstitution methods.

## Summary

It has been shown that the CF expressed CYP3A4 monooxygenase system is functional, whether expressed with CPR alone or with both CPR and C*b5*. Higher levels of activity were observed with the 2-enzyme system than the 3-enzyme CYP3A4 monooxygenase system, an effect of the increased CPR to CYP3A4 ratio, leading to increased electron transfer to CYP3A4. The crude CF expression product, containing both proteopolymersomes and non-polymersome bound functional enzymes, had the highest levels of activity. This “crude” sample was used for the purification of the proteopolymersomes. The centrifuged sample presented decreased levels of activity at around a third of the total, as expected, with the pelleting of non-polymersome bound CYP3A4. The filtration and separation of the total sample however lead to loss of all activity, due to the disturbance of protein structure and the dilution of sample respectively. Avidin purification of the (proteo)polymersomes was the most successful method of purification. All aggregates and impurities were removed, and activity levels were up to half of those seen in the crude product. The specific activity however was not as high as that seen with the commercial CYP3A4 Baculosome preparation, indicating either misfolded inactive proteins present on the proteopolymersome or possibly inefficient presence and crowding of CPR and C*b5* to optimise activity of CYP3A4.

**Chapter 6:**

# Characterisation of the CYP3A4 Proteopolymersome

6.1 INTRODUCTION

6.2 AIMS

6.3 EXPERIMENTATION

6.4 STABILITY OF THE PROTEOPOLYMERSOME

6.5 FLOW CYTOMETRY ANALYSIS

6.6 TEM IMAGING

6.7 DISCUSSION

6.8 SUMMARY

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

DLS – Dynamic light scattering

DPPE – 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine

DSPE – 1,2-distearoyl-sn-glycero-3-phosphoethanolamine

FACS – Fluorescence-activated cell sorting

FITC – Fluorescein isothiocyanate

PEG – Polyethylene glycol

PMSF – Phenylmethanesulfonyl fluoride

SEM – Standard error of the mean

TEM – Transmission electron micrograph

List of Tables

**Table 6.1:** Vesicle count of polymersomes and Baculosomes over the period of 28 days.

List of Figures

**Figure 6.1:** Monitored activity levels of CYP3A4 proteopolymersomes and CYP3A4 Baculosome over 28 days.

**Figure 6.2:** DLS analysis of polymersomes and Baculosomes over 28 days.

**Figure 6.3:** The effect of additives on the activity of CYP3A4 monooxygenase proteopolymersome.

**Figure 6.4:** Monitored activity levels of CYP3A4 proteopolymersomes and CYP3A4 Baculosome over 28 days in the presence of additives.

**Figure 6.5:** Flow cytometry analysis of PBD-PEO polymersomes.

**Figure 6.6:** Count of polymersomes showing rhodamine B fluorescence.

**Figure 6.7:** Flow cytometry analysis of PBD-PEO (proteo)polymersomes.

**Figure 6.8:** Overlap of the count of (proteo)polymersomes showing rhodamine B fluorescence.

**Figure 6.9:** Plot of rhodamine B and AlexaFluor 488 fluorescence of the (proteo)polymersomes.

**Figure 6.10:** Overlap of the (proteo)polymersomes populations showing FITC fluorescence.

**Figure 6.11:** Fluorescence readings of rhodamine B and AlexaFluor 488 fluorescence following avidin purification.

**Figure 6.12:** TEM image of osmium oxide stained PBD-PEO polymersomes.

**Figure 6.13:** TEM image of immunogold labeled proteins on proteopolymersomes.

**Figure 6.14:** TEM image of immunogold labeling of polymersomes.

## Introduction

After the development of functional and pure microsomal monooxygenase proteopolymersomes, it is important to characterise these vesicles and understand their advantages as well as their limitations and drawbacks, so they can be challenged and improved upon. One important characteristic of any proteopolymersome, and indeed one of its main advantages of over other alternative lipid membrane environment, is its stability. It is acknowledged that in the polymersome membranes the self-assembled polymers form a tight and entangled arrangement that is more stable than the lipid bilayer arrangement, and the life time of polymer vesicles is much longer than that lipid based structures (Discher et al., 2002, Discher and Ahmed, 2006). It is possible that such stability of the polymersomes and their membranes can also translate into improved protein functional life span for proteopolymersomes, and the maintenance of functionality longer than with lipid-based preparations. With proteopolymersomes maintaining their structural integrity, the anchored proteins will always have their stabilising environment and maintain their folded and active conformations. Activity testing of purified CYP3A4 monooxygenase proteopolymersome is an indicator of long term enzyme stability and conformation, while DLS analysis can monitor the structural integrity of proteopolymersome vesicles.

Although at the point of purification the proteopolymersomes are clear of any proteolytic elements that can affect long term functionality, problems can arise with prolonged storage. Microbial growth is a problematic complication that can arise with storage of *in vitro* preparations, leading to release of many undesired contaminants including detrimental proteases which can severely diminish structural integrity and functionality of proteins. Additives have been used for the long-term storage of *in vitro* preparations including biocides that target microbial growth *e.g.* metabolic enzyme function inhibitors such as sodium azide and Prolin, and antibiotics such as ampicillin and kanamycin, as well as other functional additives such protease inhibitors *e.g.* PMSF and protease inhibitor cocktails, and cryoprotectants for sub-zero temperature storage *e.g.* glycerol. These additives do have limitations and are chosen dependent on function, biocides such as sodium azide can interfere and strongly inhibit the *in vitro* prepared enzymes’ function, while the protease inhibitor PMSF for example is only active against a subset of proteases, in addition to having a short functional life span in aqueous solutions.

Beyond the stability characterisation of the proteopolymersome, the enzyme co-translational insertion and the structural integrity through the purification procedure can also be characterised. The amount of protein inserted onto the polymersome membrane is an important metric in the development of a multi-enzyme system. With large excess of polymersome vesicles present, the enzymes would become more spread out on the membranes, and too far apart for close protein-protein interactions, decreasing membrane crowding required for efficient microsomal monooxygenases. While if insufficient polymersome vesicles are used the expressed protein will be unable to anchor into the membrane and instead aggregate. Measuring a population of polymersome vesicles can be carried out using flow cytometry analysis, similar to standard cell population analysis. The fluorescent labelling of proteins anchored on polymersome can identify which vesicles contain anchored proteins, and even the utilisation of FACS can be applied for the separation of the active proteopolymersomes.

Analysis of the polymersome structural integrity can also be confirmed through the use of DLS analysis and TEM imaging. DLS can be used to track polymersome count to confirm no loss of vesicles, and confirm maintenance of monodispersity of size in the vesicle population. TEM imaging can show the vesicular structure of polymersomes is maintained, especially for the vesicles that have undergone the purification process. Immunostaining of proteins with gold nanoparticle-conjugated antibodies can confirm the presence of protein on the polymersome membrane.

## Aims

Analysing the stability of the CYP3A4 microsomal monooxygenase proteopolymersome through functionality testing over the duration of a month, followed with optimisation of activity through use of additives. The use of flow cytometry techniques to monitor polymersome shape and presence of protein following avidin purification of proteopolymersomes. Analysis of polymersome stability through monitoring of vesicle count and size via DLS measurements, and the structure through TEM imaging. The use of immunogold labelling techniques to confirm the presence of protein anchored on membrane surface.

## Experimentation

### Avidin purification of proteopolymersomes

Avidin purification of (proteo)polymersomes was carried out as described in the methodology, section 2.6. For the flow cytometry analysis, while bound to the avidin resin the biotinylated (proteo)polymersomes underwent immunolabelling with AlexaFluor 488 fluorescent antibody (anti-histidine tag) targeting CYP3A4, CPR and C*b5*. For TEM imaging, while bound to the avidin resin the biotinylated (proteo)polymersomes underwent immunogold labelling specifically targeting expressed proteins.

### Stability analysis

Avidin purified CYP3A4 monooxygenase (proteo)polymersomes were tested for activity, as described in the Methodology, section 2.7. Samples were kept at 4oC, and activity monitored at regular intervals (start, 4 days, 7 days, 14 days and 28 days). Similar storage and activity conditions were used for Baculosome sample as comparison.

### Flow cytometry

Flow cytometry analysis was carried out as described in Methodology, section 2.8. Fluorescent and non-fluorescent polymersome were run to establish polymersome size and complexity, as well as fluorescence. Fluorescent polymersomes and proteopolymersomes were then run to monitor AlexaFluor 488 fluorescence and presence of polymersome-anchored proteins.

### DLS analysis

DLS analysis was carried out as described in Methodology, section 2.8. All stages of avidin purification of (proteo)polymersomes were measured for size, polydispersity index and count. DLS measurements was also carried out on all stages of stability study to confirm stability of polymersome vesicle.

### TEM Imaging

Prepared polymersomes and CYP3A4 monooxygenase proteopolymersomes were imaged with TEM as described in Methodology, section 2.8. Presence of immunogold nanoparticles were analysed on both populations to confirm presence of protein on polymersome membrane.

## Stability of the proteopolymersome

Although starting off with various levels of overall activity, when normalised to the same level, the purified CYP3A4 microsomal monooxygenase proteopolymersome and the Baculosome have very different activity profile over the period of 28 days (figure 6.1). The Baculosome can be used as a model for microsomal monooxygenase system on a lipid membrane, and compared to the monooxygenase system on a polymer membrane. CYP3A4 activity on the Baculosome falls sharply after first day, and by fourth day nearly 60% loss of activity. After 28 days, there is almost no observable CYP3A4 Baculosome activity remaining with activity levels at just 4% of that observed at the start. Unaltered Proteopolymersome has a better activity profile, with a loss of only 30% of activity after 4 days. After 2 weeks, the decrease in activity levels of the proteopolymersomes matches that of the Baculosome after 4 days at around 60% decrease. While the proteopolymersome does have a low-level activity after 28 days with 80% loss of activity, it’s almost five times higher than seen with Baculosomes.

|  |
| --- |
|  |
| **Figure 6.1:** Monitored activity levels of CYP3A4 proteopolymersomes and CYP3A4 Baculosome over 28 days. Activity levels were monitored at 2 hours post start of reaction and normalised to activity seen at the first day. Blue: CYP3A4 monooxygenase proteopolymersome, yellow: CYP3A4 Baculosome. Error bars represent the SEM of three independently repeated experiments. |

The activity profile of the proteopolymersome although markedly better than that of the Baculosome, has room for improvement. The possible causes of this are either polymersome driven *e.g.* loss of polymersome structure and integrity or more likely due to protein degradation through protease activity or protein unfolding.

The effect of vesicle integrity on the microsomal monooxygenase activity was examined by monitoring DLS analysis of the vesicles over the period of 28 days (figure 6.2, table 6.1). Both the size profile and vesicle count can be used as indicators of the stability of the vesicles and the integrity of their membranes.

|  |
| --- |
| **a)** |
|  |
| **b)** |
|  |
| **Figure 6.2:** DLS analysis of polymersomes and Baculosomes over 28 days. **a)** Size profile of (proteo)polymersomes at 0, 14 and 28 days. **b)** Size profile of Baculosomes at 0, 14 and 28 days. Blue: at 0 days, orange: after 14 days, and grey: after 28 days. |

It is clear from the DLS data that while polymersomes maintain their size profile and vesicle count over the period of 28 days, the Baculosomes are much less stable. The Baculosomes start with three distinct population sizes that over the period of 28 days converge into a single monodisperse population. The vesicle count tells a similar story, with the count of the polymersomes remaining relatively constant over the period of 28 days, while the Baculosomes decrease significantly in count, with around 73% of vesicles remaining after 28 days. These results while they can give an explanation to the decrease of activity seen with the Baculosomes, they do not answer why there was a decrease in activity with the proteopolymersomes. The polymersome size profile and counts remains relatively constant throughout the 28 days. This shifts the attention towards multiple possible causes including enzyme stability and microbial growth in the in vitro preparations, coupled with protease production.

|  |
| --- |
| **Table 6.1:** Vesicle count of polymersomes and Baculosomes over the period of 28 days. |
| |  |  |  |  | | --- | --- | --- | --- | |  | **Start** | **14 days** | **28 days** | | **Polymersomes** | 183.1 | 191.4 | 193.6 | | **Baculosomes** | 258.6 | 191.5 | 183.8 | |

The supplementation of additives in the CYP3A4 proteopolymersome preparation can elongate the life expectancy of the sample. With the main concern being possible microbial growth causing increase in protease levels and loss of function, additives were selected to improve the long-term activity levels. Sodium azide and kanamycin were selected to inhibit bacterial growth, while PMSF and protease inhibitor tablets to inhibit protease activity. Glycerol was selected as a useful additive for enhancing CYP3A4 proteopolymersome stability following multiple freeze-thaw cycles. Firstly, the effect of each additive on CYP3A4 activity was monitored (figure 6.3).

|  |
| --- |
|  |
| **Figure 6.3:** The effect of additives on the activity of CYP3A4 monooxygenase proteopolymersome. Error bars represent the SEM of three independently repeated experiments. |

It is clear that sodium azide is an inhibitor of CYP3A4 monooxygenase activity with 80% loss of activity, and therefore unsuitable for use as an additive. PMSF, while having minimal effect on CYP3A4 proteopolymersome activity, is not an ideal additive for long term analysis due to lack of stability and its short-term activity of only a few hours. Kanamycin and the protease inhibitor cocktail are the most ideal additives for their minimal effect on activity and long active life. The cryoprotectant glycerol is also very advantageous as it has no effect on CYP3A4 activity and can help maintain proteopolymersome activity following freeze-thaw cycles.

The use of additives proved beneficial for the improvement of proteopolymersome stability (figure 6.4). While the protease inhibitor cocktail provided no detectable improvement in stability levels, most likely due to the proteases present in the preparation not being affected by the inhibitor cocktail, the use of kanamycin was successful.

|  |
| --- |
|  |
| **Figure 6.4:** Monitored activity levels of CYP3A4 proteopolymersomes and CYP3A4 Baculosome over 28 days in the presence of additives. Activity levels were monitored at 2 hours post start of reaction and normalised to activity seen at the first day. Blue: kanamycin supplemented CYP3A4 monooxygenase proteopolymersome, orange: protease inhibitor cocktail supplemented CYP3A4 proteopolymersome and grey: kanamycin supplemented CYP3A4 Baculosomes. Error bars represent the SEM of three independently repeated experiments. |

With kanamycin preventing bacterial growth and minimising any production of proteases, the maintenance of activity was superior than without the use of any additives, with over 75% of activity maintained after 28 days for kanamycin supplemented proteopolymersomes, compared to 18% with proteopolymersomes only and 4% in Baculosomes only. In fact, no loss of proteopolymersome activity was observed for the first 14 days. While even with kanamycin supplementation, the Baculosome activity profile followed a similar profile as that seen without any supplementation.

## Flow cytometry analysis

Flow cytometry analysis of (proteo)polymersomes can help elucidate the population size of polymersomes and confirm the insertion and anchoring of proteins onto the polymersome membrane. Firstly, the polymersomes, both non-fluorescent and fluorescently labelled were run through the flow cytometry (figure 6.5).

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 6.5:** Flow cytometry analysis of PBD-PEO polymersomes. **a)** non-fluorescent polymersomes. **b)** fluorescent rhodamine labelled polymersomes. | |

The size (forward scatter) and complexity (side scatter) profiles for both polymersomes are very similar, showing that the addition of extra functional groups, such as DPPE-lissamine-rhodamine B or DSPE-PEG-biotin, to polymersomes does not alter the profile significantly.

However, the fluorescence profile is significantly different (figure 6.6). The non-fluorescent vesicles, as expected, show no rhodamine B fluorescence. The labelled polymersomes show 2 distinct populations, around 40 % of the polymersomes show no or weak levels of fluorescence, while nearly 60% show high detectable levels of fluorescence. These results indicate that not all the polymersomes while extruding incorporate DPPE-rhodamine B into their membrane.

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 6.6:** Count of polymersomes showing rhodamine B fluorescence. **a)** non-fluorescent polymersomes. **b)** fluorescent rhodamine labelled polymersomes. | |

Following the analysis of polymersomes, the (proteo)polymersomes were run through the flow cytometry. The two (proteo)polymersome samples, one with histidine tag CYP3A4 and CPR and one without, are present on rhodamine B labelled fluorescent polymersomes. The two samples underwent immunolabelling with polyhistidine tag specific AlexaFluor 488 antibodies. The size and complexity profile of these two samples (figure 6.7) was similar to that seen of standard polymersome samples in figure 6.5. Over half the population of detected particles was present in a small area of the plot, which was gated (population 1), and selected for further analysis. The population 1 rhodamine B fluorescence profile of both (proteo)polymersome populations are very similar (figure 6.8). The same distinct “2 population” profile as seen with polymersomes (figure 6.6b) is observed. The rhodamine B fluorescence populations, are known to be polymersomes, and were gated (population 2) for further analysis of antibody binding to membrane bound proteins.

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 6.7:** Flow cytometry analysis of PBD-PEO (proteo)polymersomes. **a)** Avidin purified polymersome sample. **b)** Avidin purified CYP3A4 monooxygenase (proteo)polymersome. | |

Population 2, for both the sample containing protein and the sample without, are analysed for FITC fluorescence (of the AlexaFluor 488 antibody). While they share similar rhodamine B fluorescence profile, there is a distinct difference in the FITC fluorescence profile. The purified polymersome only sample (figure 6.9a) show a majority of FITC fluorescence values of below 100. With the use of this value as a threshold for AlexaFluor 488 binding and increased FITC fluorescence, the increase in FITC fluorescence for the CYP3A4 monooxygenase (proteo)polymersome can be identified and quantified.

|  |
| --- |
| **2** |
| **Figure 6.8:** Overlap of the count of (proteo)polymersomes showing rhodamine B fluorescence. **Blue:** Avidin purified polymersome sample. **Red:** Avidin purified CYP3A4 monooxygenase (proteo)polymersome. |

While the empty polymersome sample presents with only 3.6% of the population as being FITC positive and 96.4% as being FITC negative, the CYP3A4 monooxygenase proteopolymersomes present with a markedly higher 27.3% of the population as FITC positive and 72.7% of the population as FITC negative (figure 6.9).

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 6.9:** Plot of AlexaFluor 488 (protein tagging) fluorescence of the (proteo)polymersomes. FITC fluorescence is the fluorescence of the AlexaFluor 488 antibody, PE-A fluorescence is the fluorescence of the rhodamine B. **a)** Avidin purified empty polymersome sample. **b)** Avidin purified CYP3A4 monooxygenase (proteo)polymersome. | |

Overlapping the populations can clearly visualise the increase in FITC fluorescence observed on the CYP3A4 monooxygenase proteopolymersome (figure 6.10). This was also confirmed by the fluorescence reading of the avidin purification elutes (figure 6.11). While the fluorescent polymersomes showed high levels of rhodamine B fluorescence following purification, only those with expressed CYP3A4 monooxygenase enzymes presented with FITC fluorescence.

|  |
| --- |
|  |
| **Figure 6.10:** Overlap of the (proteo)polymersomes populations showing FITC fluorescence. The count of (proteo)polymersomes showing AlexaFluor 488 fluorescence of the empty polymersome population (red), and CYP3A4 monooxygenase (proteo)polymersome (blue). |

|  |
| --- |
|  |
| **Figure 6.11:** Fluorescence readings of rhodamine B and AlexaFluor 488 fluorescence following avidin purification. Blue: rhodamine B fluorescent CYP3A4 proteopolymersomes, orange: rhodamine B fluorescent polymersomes, and grey: non-fluorescent polymersomes. |

## TEM Imaging

TEM imaging of (proteo)polymersomes can help display their structure and confirm its stability following the purification processes. Maintenance of the vesicular structure of polymersomes is important for the stability of any bound protein, to continue providing the stabilising membrane environment. This vesicular structure of polymersomes was observed for standard PBD-PEO diblock polymersomes (figure 6.12). While the majority of the polymersomes present with the typical vesicular structure, there are a few presenting with structural deformities, possibly due to the dry preparation of the samples during staining and imaging. This is also coupled with apparent shrinkage of polymersome, with vesicle diameters in the range of 50 – 100 nm, rather than the expected 100 – 200 nm range seen previously in DLS measurements.

The immunogold labelling of proteins on CYP3A4 monooxygenase (proteo)polymersomes can confirm the presence of proteins bound to the membrane. The gold nanoparticle conjugated antibodies target the polyhistidine tags present only on the CYP3A4, CPR and C*b5* proteins. The gold nanoparticles appearing as dark spots on the TEM image can be quantified to show the number of proteins present on each polymersome. The imaging of immunogold labelled (proteo)polymersomes showed distinct presence of gold nanoparticles present on the membrane of the CYP3A4 monooxygenase proteopolymersomes (figure 6.13), while no gold nanoparticles were observed for the polymersomes (figure 6.14).

|  |
| --- |
|  |
| **Figure 6.12:** TEM image of osmium oxide stained PBD-PEO diblock polymersomes. Bar represents 200 nm. |

|  |
| --- |
|  |
| **Figure 6.13:** TEM image of immunogold labeled proteins on proteopolymersomes. Bar represent 100 nm. |

|  |
| --- |
|  |
| **Figure 6.14:** TEM image of immunogold labeling of polymersomes. Bars represent 100 nm. |

The speckled nature of proteopolymersomes, in figure 6.13, indicate that proteins are wide spread across the polymersomes and do not assemble into certain regions in close proximity to each other. The polymersome only sample does not show any gold nanoparticle speckling.

## Discussion

The stability of CYP3A4 monooxygenase proteins is superior when present on the membrane of polymersomes than when present on lipid membranes such as the microsomes. This is translated to increased retention of CYP3A4 activity. While CYP3A4 present on the lipid membrane of the Baculosome loses almost all its activity after 28 days, CYP3A4 present on the polymersome retains almost 20% of the activity. This increased stability is due to multiple reasons, one is the purity of the sample following purification. Without any impurities, protein or otherwise, and minimal microbial presence there is much less protein degradation, which aids in the maintenance of protein structure and folding. The structure and organisation of polymersomes also help greatly in driving the stability of the CYP3A4 monooxygenase system. The compact and strong arrangement of the diblock copolymers means that the polymersomes remain as vesicles far longer than their lipid counterparts, as shown with the DLS analysis. The constant presence of polymersomes allows the proteins to always have the stabilising membrane environment required for enzyme folding and function.

Although having superior stability than lipid-based structures, there is still a lot of protein degradation and the stability of the proteins bound to polymersomes can still be further improved. The use of additives to the purified proteopolymersomes were examined and showed varying effects on the monooxygenase activity of CYP3A4. The bactericidal sodium azide diminishes most of CYP3A4’s activity, while the antibiotic kanamycin showed minimal effect on the enzyme’s activity. The use of these antimicrobial agents is key in preventing bacterial growth that can produce harmful proteases that cause breakdown of the enzymes. Currently, the storage of the proteopolymersomes at 4oC alone is insufficient at preventing and minimising microbial growth. Additionally, the use of proteases was tested, and both PMSF and protease cocktail mix showed no significant effect on CYP3A4 activity. However, though the suitability of these two different proteases for stability purposes varies greatly. PMSF is itself only stable for a few hours and is only specific against most members of one class of proteases, the serine proteases. This limited active life span and limited inhibition of PMSF limits its usefulness as an additive for proteopolymersome preparations. The protease cocktail mix has several advantages over PMSF, including the broader inhibition of multiple classes of proteases, and the longer life expectancy and irreversibility. The action of both protease inhibitors and antimicrobials can greatly improve the stability of the proteopolymersomes.

This was observed with the use of kanamycin, which is more stable than other antibacterial agents such as ampicillin. The stability conferred by the presence of kanamycin was clearly observed in the activity of the microsomal monooxygenase proteopolymersome, with the maintenance of all activity after 2 weeks, and while the proteopolymersome alone maintained less than a fifth of the original activity after 28 days, three-quarters of the activity was maintained after 28 days when kanamycin prevented bacterial growth. Interestingly, for the first 7 days of storage, the proteopolymersomes in the presence of kanamycin perform better than at the start, with a 20% increase in activity levels, and its only after 14 days that activity levels return to the same as those seen at the start. This boost in activity can perhaps be attributed to the extra time required for protein folding and incorporation of cofactors while present on the polymersome membrane, or possibly the time required for protein movement across the membrane for successful protein-protein interactions. The lack of improvement seen in the Baculosomes supplemented with kanamycin highlight the inherent problem of instability present with lipid based membrane structures, and one which elevates the importance of more stable polymer based membranes.

Flow cytometry analysis of the proteopolymersomes has confirmed the presence of proteins, CYP3A4, CPR and C*b5* on the membranes. The size and complexity scatter show that the avidin purified (proteo)polymersome sample has high levels of uniformity, with the majority of the measured vesicles in the population occupying a very small range of values, similar to that of the polymersome only samples, and showing that they are unaffected by the avidin purification processes. Contrasting that, in the presence of impurities such as the wheat germ CF components, the scatter would have given a large distribution of particles across a wide range of size and scatter.

The rhodamine B fluorescence profile of the fluorescently labelled (proteo)polymersomes is an interesting find of this data as it highlights the fact that not all polymersomes take up and incorporate lipid bound moieties into their membranes such as DPPE-lissamine-Rhodamine B. The biotin tag used for biotinylating of polymersomes is also lipid linked as DSPE-PEG-biotin. If the same trend seen with the DPPE lipid uptake were to be observed with DSPE lipid, that would indicate that a minority of polymersomes do not get biotinylated via DSPE-PEG-Biotin, and are lost during the avidin purification process. This could be the main cause for the loss of unbound polymersomes, nearly 40% of total polymersomes, during the avidin purification process seen in figure 4.8, as these polymersomes were fluorescently labelled but however not biotinylated while those on the beads were both fluorescent and biotinylated. Minimising the loss of these vesicles is key in optimising the yield of proteopolymersomes as many of them would have bound proteins on their membranes. Caution must be taken when increasing DSPE-PEG-biotin concertation during polymersome synthesis process to improve the number of vesicles that are biotinylated as the increase in biotin concentration could complicate the vesicle elution off the avidin resin and require higher biotin concentrations to obtain similar yields.

The AlexaFluor 488 antibody fluorescence tells a lot about the (proteo)polymersomes too. These antibodies bind only to the expressed proteins, and as can be seen in figure 6.9, there is a marked increase in the fluorescence. The fluorescence difference between the protein and no protein samples indicate that almost 25% of the polymersomes contain membrane bound proteins, as shown by the increase in fluorescence. Although the polymersomes with fluorescence levels below the fluorescence marker levels are not necessarily without bound proteins, but possibly the proteins do not have bound antibodies, or the number of proteins present on the membrane are so low that the fluorescence of the AlexaFluor 488 antibody on them is undetected. One key take away from this is that the polymersomes added to the CF expression mixture, 30 μl of 10 mg/ml PBD-PEO for wheat germ based expression, is in excess to the amount required for the actual amount of protein expressed. Optimisation of polymersome supplementation for a multi-enzyme system is very important. Minimising the presence of unused vesicles will ensure the proteins that do anchor into vesicles that have a high number of proteins, increasing the crowding and chance of enzyme interaction, electron transfer and a functional monooxygenase. However, minimising the supplementation of vesicles too much will firstly decrease the chance of a protein interacting with it and inserting into its membrane, and if there is no available free membrane to insert into then it will drive the formation of undesired aggregates.

The TEM images, figure 6.13 and 6.14, confirm what was seen with the flow cytometry, that the avidin purified (proteo)polymersomes are similar in shape and size as those of fresh polymersomes. This confirms that polymersomes are unaffected by the avidin purification process and can maintain their vesicular structure. The gold immunolabelling of proteopolymersomes give similar findings to the flow cytometry analysis, with only a fraction of the polymersomes show immunogold labelling while others do not. A very small amount of (non-protein bound) polymersomes showed signs of gold immunolabelling although that is likely due to non-specific binding of the antibody to the polymers since at most these vesicles show one or two gold nanoparticles, compared to the high number of gold nanoparticles seen on the proteopolymersomes.

There are some visible structural changes for the polymersomes, as well as size alteration. Many of the polymersomes seem to have average diameter of less than 100 nm, while most DLS readings show that the majority of polymersomes are over 100 nm. This variation in shape and size can be attributed to the TEM imaging film preparation process, including the drying, fixing and staining of the (proteo)polymersome samples. Similar problems have been encountered by other groups (Zapf et al., 2015) and has been countered by the use of glutaraldehyde fixation and production of polymersomes with trehalose solution, a known preserver of proteins and membranes (Christensen et al., 2008).

## Summary

The characterisation of the CYP3A4 monooxygenase proteopolymersome has elucidated the properties of both the proteins and the polymersome. The analysis of the stability of the proteopolymersome has shown that while polymersomes offer far superior stability than lipid based vesicles, there is still protein degradation that needs to be prevented either by the inhibition of microbial growth or the inhibition of proteases. The flow cytometry has confirmed the presence of protein bound to the polymersome membrane and shown the percentage of polymersomes in a sample that have proteins embedded on the membrane, at roughly 25%. Similarly, the TEM imaging confirms the presence of membrane bound proteins and that the structure of polymersomes is maintained from the protein expression stage through to the purified preparation stage.

**Chapter 7:**

# Development of Peptide-Based Polymersome Anchors

7.1 INTRODUCTION

7.2 AIMS

7.3 EXPERIMENTATION

7.4 EXPRESSION AND PURIFICATION

7.5 RECONSTITUTION AND SEC

7.6 FLUORESCENCE MICROSCOPY

7.7 DISCUSSION

7.8 SUMMARY

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

EGFP – Enhanced green fluorescent protein

GUV – Giant unilamellar vesicle

HLT – Histidine tag lymphotactin tev protease

PDMS – Polydimethylsiloxane

PMOXA – Poly-2-methyl-2-oxazoline

POPC – 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

SEC – Size exclusion chromatography

List of Figures

**Figure 7.1:** Surface decoration of polymersomes using peptide anchors.

**Figure 7.2:** SDS-PAGE of *E. coli* expressed EGFP fusion peptides.

**Figure 7.3:** SDS-PAGE of *E. coli* expressed soluble HLT-EGFP fusion protein.

**Figure 7.4:** SEC separation profile for soluble EGFP and melittin:EGFP fusion.

**Figure 7.5:** SEC separation profile for soluble EGFP and polymersomes.

**Figure 7.6:** SEC separation profile for melittin:EGFP fusion protein with and without Brij detergent present.

**Figure 7.7:** SEC separation profile for melittin:EGFP fusion protein before and after reconstitution with polymersomes.

**Figure 7.8:** SEC separation profile for polymersomes, before and after reconstitution.

**Figure 7.9:** Fluorescence microscopy of electroformed rhodamine B labelled liposome GUVs reconstituted with melittin:EGFP fusion protein.

**Figure 7.10:** Time lapse capture of EGFP encapsulated liposome GUVs, following the addition of melittin peptide.

**Figure 7.11:** EncapsulatedEGFP fluorescence intensity.

**Figure 7.12:** Fluorescence microscopy of melittin:EGFP reconstituted onto triblock polymersomes.

**Figure 7.13:** EGFP fluorescence readings of CF expressed EGFP and melittin:EGFP.

## Introduction

Polymersomes on their own are typically non-reactive, and their functionalisation requires further modifications of their surface for specialised applications. Typically, lipid bound moieties have been successfully used for the surface decorations of polymersomes, with the lipid portion inserting into the membrane, while the functional group is exposed to the external environment. The use of lipid-based moieties is limited in application and involves complex synthesis and purification protocols. Alternatively, membrane binding peptides have been used as a replacement for the lipid moieties (figure 7.1). The fusion of proteins to the membrane binding peptides can be done with simple cloning techniques. The binding capabilities of various peptides to polymersomes has already been examined, including the antimicrobial peptide Cecropin A, and the membrane binding region of C*b5* (Noor et al., 2012, Klermund et al., 2016). These peptides have been examined for their polymersome membrane binding ability, and their effect on polymersome stability.

|  |
| --- |
|  |
| **Figure 7.1:** Surface decoration of polymersomes using peptide anchors. The peptide anchors (red) can be fused to a wide range of soluble, and non- membrane binding proteins. |

In this chapter the polymersome membrane binding abilities of three antimicrobial peptides; cecropin A, melittin, and LAH4 will be examined. These three peptides have been fused to a C-terminus polyhistidine tagged EGFP, to be used as a marker for successful insertion into the polymersome membrane. The anti-microbial peptides chosen here have key characteristics that enables and drives their insertion into lipid membranes. Cecropin A, of insect cell origin, has a helical structure that targets membranes and in high numbers, and can form pores and ion channels that destabilise membranes and lyse cells (Sato and Feix, 2006). While such a peptide can cause major problems with lipid vesicles, the increased structural integrity and stability of polymersomes should be better prepared to cope with destabilising effects of the peptide, especially more so with a soluble protein fused to it. Melittin, similar in structure to cecropin A, is also a short helical peptide of insect origin. The peptide is itself amphiphilic, with a hydrophobic N-terminus and a hydrophilic C-terminus (Terwilliger and Eisenberg, 1982, Bechinger, 1997). There are multiple ways in which melittin targets cellular lysis, including the inhibition of transport pumps and the disruption of negatively charged lipids and pore formation leading to increased membrane permeability (Sato and Feix, 2006, Leveritt et al., 2015). LAH4 is also amphipathic and anti-microbial as melittin, however this designer peptide possesses other unique and interesting characteristics (Bechinger, 1996, Perrone et al., 2014). Its structure of repeated lysine, alanine and histidine residues give it and other peptides derived from it unique properties not present in cecropin A or melittin. Its interactions with the membrane are strong and have been used for both anti-microbial purposes as well as for transfection and cell penetrative purposes (Moulay et al., 2017). It has been shown to be able to act as a carrier into cells carrying molecules and enzymes bound to it covalently across the membrane. Due to its histidine rich nature, its interactions with lipid membrane have been shown to be affected by its environment (Georgescu et al., 2010), both pH and temperature are known to affect the orientation, insertion and anchoring into the membrane.

The effect of these antimicrobial peptides and their polymer membrane binding and inserting abilities has not been examined in detail. So far the melittin peptide has been shown to cause the rupture of liposomes, but not of polymersomes (Vijayan et al., 2005). However, the analysis of melittin binding to the membrane, either as an individual peptide or part of a fusion protein has not been carried out. Cecropin A has been shown to interact and stably insert into triblock polymersome via reconstitution methods (Noor et al., 2012). LAH4 is untested with polymersomes, both for its effect on polymersome structure and stability, and its membrane binding ability.

Most of the research carried out so far on polymersome peptide anchors has been done through the use of cell-based expression followed by purification and reconstitution. Similarly, to start off with, the three fusion peptides will be expressed using *E. coli* cell-based expression system, followed by affinity purification. Analysis of the membrane binding ability of the peptides will be compared for polymersomes and liposomes. Peptide reconstitution methods can be applied to the vesicles and examined under a fluorescence microscope, looking for EGFP fluorescent membranes. Additionally, the use of SEC separation following reconstitution onto the vesicle membranes can highlight the separation profile of free protein, and that bound on the membrane. Polymersome supplemented CF systems can also be used for the development of peptide anchored proteopolymersomes. Analysis of the peptide co-translational insertion into the polymersome membrane can be proven by fluorescence analysis of the purified proteopolymersome, flow cytometry, and fluorescence microscopy. Further steps can be taken in the development of these peptide anchors, the use of protein engineering approaches such as directed evolution and rational design can alter the characteristics further into more desired traits such as improved hydrophobicity and membrane binding, or increased histidine residues to control peptide orientation.

Establishing the polymer membrane binding abilities of these peptides open the avenues for the functionalisation of polymersome surface with a wide range of soluble proteins, without the need for complex protein-polymer conjugations. The immobilisation of enzyme can offer can benefits for bio-catalysis offering improved enzyme activity and stability. As well these peptide anchors can be used for the anchoring of insoluble membrane proteins that have difficulty inserting themselves into polymersome membranes. The replacement of the hydrophobic membrane binding regions with uniform and carefully designed peptides with known polymersome binding capacities can simplify the insertion of many membrane proteins and as such their study in their native environment is also simplified.

## Aims

The cloning, expression and purification of three anti-microbial peptides fused to EGFP. Analysis of their membrane binding capabilities via reconstitution methods through the use of SEC and fluorescence microscopy, and their co-translational insertion capabilities through CF expression. The use of membrane binding assays to examine the co-translational insertion of the anti-microbial peptides and the identification of ideal peptide sequences for polymersome binding. The engineering of these sequences can develop peptide anchors with strong polymer membrane interactions and anchoring capacity.

## Experimentation

### Cloning, expression and purification

Cecropin A, Melittin and LAH4 peptide sequences were cloned into pET24a vector for *E. coli* expression, as stated in the Methodology section 2.2. The peptides were fused to the c-terminus of the *egfp* gene using a polyalanine (x10) linker, while a polyhistidine tag was added at the n-terminus of *egfp*. Expression was carried out in Rosetta *E. coli* cells under the control of a T7 promoter, as stated in the Methodology section 2.4. Transformed cells were grown at 37oC till OD 0.7, followed by IPTG induction of expression, and incubation at 30oC for 18 hours. *E. coli* cells were harvested and sonicated for cell lysis. Polyhistidine tag-nickel purification was carried out on the fusion peptides using HisTrap HP 5 ml column, as stated in the Methodology section 2.6. EGFP and melittin:EGFP were expressed using the wheat germ CF protein expression system. Biotinylated polymersome supplementation was carried out as described in the methodology, section 2.4. The EGFP fluorescence levels was measured for each protein following expression and following avidin purification, as stated in the Methodology section 2.6.

### Reconstitution and SEC

The concentration of the fusion peptides was calculated and were diluted to 1 mg/ml. The peptides were incubated with fluorescently labelled PDMS-PMOXA-PDMS triblock polymersomes at 4oC for three hours, with shaking at 700 rpm. For SEC, the protein only, polymersome only, and the reconstituted proteopolymersomes samples were run through a Superdex 200 5/150 GL column, and all separation fractions were collected for fluorescence analysis, as described in the Methodology section 2.6.

### Fluorescence microscopy

PDMS-PMOXA-PDMS polymersomes reconstituted with the three peptide fusions were imaged using Olympus BX51 microscope, as stated in the Methodology section 2.8. Rhodamine B labelled liposome GUVs were produced using the electroformation method, with and without the presence of the fusion peptides using Olympus BX51 microscope, as stated in the Methodology section 2.8.3.

## Expression and purification

The three EGFP fusion peptides, cecropin A, melittin and LAH4 were successfully expressed and affinity purified (figure 7.2). Similarly, the soluble and polyhistidine tagged EGFP (HLT-EGFP) was expressed and affinity purified (figure 7.3). The overall expression levels of the three EGFP fusion peptides varies from peptide to peptide. Cecropin A and melittin, the native insect cell antimicrobial peptides, seemingly are readily over-expressed in E. coli cells, however the designed antimicrobial peptide LAH4 is less readily expressed. Following affinity purification, there are a few impurities bands for all the three peptides. The cecropin A fusion presents with a couple of undesired protein bands, yet the more difficult to express LAH4 and melittin present with a large number of protein band impurities. Conversely, and as expected, the soluble EGFP was very readily over expressed, and the its affinity purification yielded very highly pure and concentrated protein sample with no visible impurities. The difference in the size (kDa) between the soluble HLT-EGFP and the fusion EGFP constructs is due to the presence of a larger peptide sequence upstream of the EGFP in the HLT-EGFP protein, composed of a TEV protease site, a polyhistidine tag and a lymphotactin domain on the soluble HLT-EGFP protein (the HLT peptide sequence alone has a molecular weight of 11.9 kDa), while the peptide sequences of cecropin A, melittin and LAH4 are by comparison smaller (the molecular weights of the peptides range from 2-4 kDa).

|  |
| --- |
| **Figure 7.2:** SDS-PAGE of *E. coli* expressed EGFP fusion peptides. The three proteins were successfully expressed and purified from *E. coli* cells. Lanes 1-3: Total protein in Rosetta *E. coli* cells expressing cecropin A, LAH4, and melittin, respectively. Lanes 4-6: Nickel affinity purified cecropin A, LAH4, and melittin EGFP fusion peptides, respectively.  CecA:EGFP: 32.5 kDa, LAH4:EGFP 32 kDa, and Mel:EGFP 32 kDa. |
|  |
| **Figure 7.3:** SDS-PAGE of *E. coli* expressed soluble HLT-EGFP fusion protein. EGFP was successfully expressed and highly purified from *E. coli* cells. Lane 1: Total protein in Rosetta *E. coli* cells. Lane 2: Nickel affinity purified HLT-EGFP fusion protein.  HLT:EGFP: 41.5 kDa. |

## Reconstitution and SEC

The purified proteins, both the soluble EGFP and the fusion melittin:EGFP, were run through the SEC column. The collected fractions were measured for EGFP fluorescence to get a profile for their separations (Figure 7.4).

The separation profile of these two proteins are very different from each other. While the soluble EGFP separates into a single, and relatively narrow peak at 2.25 ml, the melittin:EGFP fusion separates into two distinct peaks, at 1 ml and 2.25 ml. This difference in separation profile indicates that the melittin peptide alters the properties of the EGFP protein significantly hence the separation into two populations. Since the new peak arises earlier than that seen for the soluble EGFP, it is larger in size and most probably a result of aggregated melittin:EGFP molecules, as seen in the SDS-PAGE (figure 7.2, lane 6), while the peak at 2.25 ml is the soluble protein fraction.

The incubation of the soluble HLT-EGFP with the polymersomes and subsequent running through SEC yielded two independent separation profiles with minimal fluorescence overlap (figure 7.5). There is no evidence of any EGFP reconstitution onto the polymersome membrane, with the large polymersomes separating out first at 1 ml, and the smaller and soluble EGFP once again separating out at around 2.25 ml.

|  |
| --- |
|  |
| **Figure 7.4:** SEC separation profile for soluble EGFP and melittin:EGFP fusion. Blue: soluble HLT-EGFP, orange: melittin:EGFP fusion. |

|  |
| --- |
|  |
| **Figure 7.5:** SEC separation profile for soluble EGFP and polymersomes. Blue: soluble HLT-EGFP, orange: triblock polymersomes. |

One problem arising from the SEC separation analysis is the overlap of the polymersome peaks with the insoluble and aggregated melittin:EGFP peak at 1 ml. The presence of EGFP fluorescence already at the polymersome peak would mask any EGFP fluorescence caused by the reconstitution of melittin:EGFP into the polymersomes. To tackle this problem, detergent was added to the affinity purified fusion protein to try and dissipate the EGFP fluorescence peak present at 1 ml separation fraction. The detergent Brij was added to the proteins, which were incubated for three hours and then run on the SEC column with the EGFP fluorescence measured to get a separation profile (figure 7.6).

|  |
| --- |
|  |
| **Figure 7.6:** SEC separation profile for melittin:EGFP fusion protein with and without Brij detergent present. Grey: melittin:EGFP fusion protein separation profile, blue: melittin:EGFP fusion protein with 0.05% v/v Brij separation profile, orange: melittin:EGFP fusion protein with 0.1% v/v Brij separation profile. |

It is clear that the addition of detergent significantly decreases the insoluble EGFP fluorescence levels at 1 ml, although not completely, while also increasing EGFP fluorescence levels at the soluble fraction at 2.25 ml. There is also a new peak arising with the use of Brij detergent, at 1.75 ml, that shows higher levels of EGFP fluorescence with increased Brij concentration. Incidentally, the higher the peak at 1.75 ml, the lower the EGFP fluorescence peak at 2.25 ml, indicating that while a low concentration of Brij at 0.05% v/v can solubilise aggregated proteins, higher concertation can acquire and lock up soluble proteins in undesired micelles.

Reconstituting the melittin:EGFP fusion protein with the triblock polymersomes, followed by dialysis to remove excess Brij detergent has caused an increase in EGFP fluorescence present in the polymersome separation fraction at 1 ml (figure 7.7). Coupled with the increase in fluorescence at the polymersome separation range of 1 ml, is the decrease in EGFP fluorescence in the micelle separation range of 1.5-2 ml, indicating that the reconstitution followed by dialysis of the detergent drives the removal of melittin:EGFP off the micelles and into the polymersome membrane. Interestingly, there is also an increase in EGFP fluorescence at the soluble protein separation range at 2.5 ml, mostly driven by the decrease in detergent availability, leading to breakdown of micelles into individual soluble proteins.

Looking at the fluorescence profile of the polymersomes, the separation of the vesicles as expected remains similar (figure 7.8). There is a slight increase in the rhodamine B fluorescence, from 1.1 ml to 2.5 ml, of the polymersomes that have reconstituted melittin:EGFP compared to the standard polymersomes only sample. This increase in fluorescence can be attributed to the increase of polymersome numbers as well as the possible transfer of rhodamine dye from the polymersomes onto the remaining detergent.

|  |
| --- |
|  |
| **Figure 7.7:** SEC separation profile for melittin:EGFP fusion protein before (blue) and after (orange) reconstitution with polymersomes. Blue: melittin:EGFP fusion protein separation profile, orange: melittin:EGFP fusion protein reconstituted with polymersomes separation profile. There is increased reconstitution of melittin:EGFP onto polymersome membranes seen with increased fluorescence at 1-1.5 ml separation range. This is coupled with decrease in melittin:EGFP fluorescence in the 1.5-2.25 ml separation range. |

|  |
| --- |
|  |
| **Figure 7.8:** SEC separation profile for polymersomes, before (blue) and after (orange) reconstitution. Blue: polymersome separation profile before reconstitution, orange: polymersome separation profile after reconstitution with melittin:EGFP. The separation profile for polymersomes remains constant before and after reconstitution. |

## Fluorescence microscopy

With the electroformation of liposomes it is possible to reconstitute proteins as the vesicles are being formed. The reconstitution of melittin:EGFP onto rhodamine labelled POPC liposomes was carried out to observe the anchoring of EGFP into the membrane (figure 7.9).

|  |
| --- |
|  |
| **Figure 7.9:** Fluorescence microscopy of electroformed rhodamine B labelled liposome GUVs reconstituted with melittin:EGFP fusion protein. The image taken showing EGFP only fluorescence, rhodamine B only fluorescence, and combined EGFP and rhodamine B fluorescence. |

It is clear that there is EGFP fluorescence, from the melittin:EGFP fusion protein, concentrated at the liposome GUV membrane. The presence of liposomes confirmed by the presence of the rhodamine B fluorescence also seen at the membrane. The lack of EGFP fluorescence within the liposomes confirms that the melittin does reconstitute on the lipid membrane, rather than remain in solution whether as aggregates or soluble.

The effect of melittin on liposomal GUV membranes was also analysed using EGFP loaded liposome vesicles. While the monitoring of EGFP fluorescence within liposomes showed gradual decrease over time due to standard fluorescence bleaching, in the presence of melittin, the decrease in fluorescence was much sharper indicating membrane pore formation and the leakage of EGFP (figure 7.10).

|  |
| --- |
|  |
| **Figure 7.10:** Time lapse capture of EGFP encapsulated liposome GUVs, following the addition of melittin peptide. |

The EGFP fluorescence with the GUVs was monitored and plotted to confirm the significant alteration in fluorescence decrease between samples without melittin and those with melittin added (figure 7.11). The extruded polymersomes at 100-200 nm are in general much smaller in size compared to liposome GUVs at 10-50 μm. The fluorescence microscopy of the melittin:EGFP reconstituted triblock polymersomes was attempted, however the vesicles were not clearly visible (figure 7.12).

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 7.11:** EncapsulatedEGFP fluorescence intensity. **a:** EGFP fluorescence levels for liposome GUVs. **b:** EGFP fluorescence levels for liposome GUVs in the presence of 1μM melittin. | |

|  |
| --- |
|  |
| **Figure 7.12:** Fluorescence microscopy of melittin:EGFP reconstituted onto triblock polymersomes. |

## Co-translational insertion

EGFP and the fusion protein melittin:EGFP were both successfully expressed using the wheat germ CF expression system. The Fluorescence of EGFP was detected for both expressed proteins crude CF expression products (figure 7.13a). The fluorescence intensity was slightly higher with EGFP compared to melittin:EGFP, as expected with the soluble EGFP protein more likely to fold in an active and fluorescent conformation, compared to the hydrophobic melittin peptide containing fusion protein which can form non-fluorescent aggregates. Following avidin purification of the (proteo)polymersomes, there is a marked difference in fluorescence levels (figure 7.13b). The polymersomes supplemented in the EGFP expression show none of the fluorescence seen in the crude CF product, while the polymersomes supplemented in the melittin:EGFP expression maintained almost 40% of the crude fluorescence, confirming co-translational insertion of melittin and the anchoring of soluble proteins to polymersomes.

|  |  |  |  |
| --- | --- | --- | --- |
| **a)** |  | **b)** |  |
| **Figure 7.13:** EGFP fluorescence readings of CF expressed EGFP and melittin:EGFP. **a)** EGFP fluorescence in crude CF expression product for EGFP and melittin:EGFP. **b)** EGFP fluorescence in avidin purified (proteo)polymersomes. Error bars represent the SEM of three independently repeated experiments. | | | |

## Discussion

It is clear that *E. coli* cells have no difficulty in expressing antimicrobial peptides, with no effect on cellular viability. However, these fusion proteins do aggregate and it complicates their purification. Cecropin A is easier to express and purify than the other two peptides, while LAH4, the designer antimicrobial peptide, is the most difficult to express in high quantity and isolate at high purity. The SEC separation profile in figure 7.3 confirms what is seen in the SDS-PAGE of figure 7.2, with two visible separation peaks for melittin:EGFP fusion protein. The first peak is that of the larger protein aggregates, while the later peak is that of soluble melittin:EGFP, and it overlaps well with the soluble EGFP (HLT-EGFP) SEC separation profile. To confirm the reconstitution of melittin:EGFP, as well as the peptide fusions, onto the polymersomes there needs to be a shift of EGFP fluorescence from its current separation profile onto an earlier peak as the EGFP is now present on the larger polymersome vesicles. This is not as easy as expected, as shown in figure 7.5, the polymersome themselves peak at the same separation fraction as the melittin:EGFP aggregates. So, if any of the fusion proteins that have gone from either the aggregates or the soluble fraction and reconstituted onto the polymersome membrane, their fluorescence would be masked by the EGFP fluorescence already there from the aggregates. To tackle the problem of the aggregate EGFP fluorescence the detergent Brij was added to help break up the aggregates and remove the EGFP fluorescence from the 1-1.5 ml range. The detergent did indeed remove the majority of the aggregated melittin:EGFP and increased the EGFP fluorescence seen in the soluble range, but a small amount of fluorescent aggregates remained, and addition of extra detergent lead to the formation of a new separation peak at 1.5 ml, seen in figure 7.6, compromised of melittin:EGFP present within detergent micelles.

Following melittin:EGFP reconstitution onto the triblock polymersomes and the removal of detergent by dialysis, there is an increase in EGFP fluorescence observed at the polymersomes separation range, as seen in figure 7.7. This is caused by the melittin reconstituting and binding on the polymersome membrane.

Similarly, melittin:EGFP was shown to reconstitute on the membrane of liposome GUVs. Rather than use detergents and SEC, electroformation of liposomes with the proteins, as shown in figure 2.2, produced vesicle sufficiently large enough to be visualised using fluorescence microscopy. Melittin:EGFP reconstitution on the membrane is clear and shown by the concentrated EGFP fluorescence present on the liposome membranes seen in figure 7.9, while little fluorescence is seen within the vesicles themselves. However, the binding of melittin on the membrane of liposomes drives the leakage of intravesicular contents. When liposomes are loaded with EGFP and captured over a time lapse, there is minimal decrease in fluorescence intensity mainly due to the EGFP fluorescence bleaching. When melittin is added the decrease in fluorescence intensity is more sharp and noticeable, as seen in figure 7.10. Figure 7.11 shows that liposomes encapsulated with EGFP lose about 25% of their intensity after 6 seconds, while in the presence of melittin, vesicles lose 50% of the fluorescence intensity after 2.5 seconds, and after 6 seconds over 70% of the fluorescence intensity is lost. This decrease in fluorescence intensity is caused by the pore forming abilities of melittin. When present in large concentrations melittin is known to assemble into leakage pores that can cause lytic breakdown of cells (Lee et al., 2013b). With liposomes, it is clear that melittin does interact and bind to the liposomes, they do not however cause the rupture of liposomes at the tested concentrations.

The effect of melittin on liposome is detrimental to its functionalisation applications due to its compromising effect on structural integrity and pore formation. Polymersomes by comparison are more rigid and compact in arrangement with a wider membrane thickness, which should mean that while melittin and other antimicrobial peptides do bind, they would not cause leakage or rupture. The problem with confirming this with EGFP loaded polymersomes is their size, at 100-200 nm it is too small to be visualised using a fluorescence microscope, so alternatives methods are required. Calcein leakage experiments have been previously used to monitor the melittin effect on polymersomes (Vijayan et al., 2005), and this similarly can be used to monitor the effect of melittin and other antimicrobial fusion peptides on the polymersomes once they have bound to the membrane. Calcein dye present within polymersomes is kept at high self-quenching concentrations, and only after dilution *e.g.* that caused by the leakage and the release of the dye into the solution does the fluorescence intensity increase significantly. Additionally, the co-translational insertion of anti-microbial fusion proteins can be examined by their CF expression in polymersomes supplemented systems, followed by purification, *e.g.* through the use of avidin purification of biotinylated polymersomes, and the analysis of their fluorescence intensities.

CF expression systems offer great advantages to the study of proteins, and for peptide anchors they offer an opportunity to examine the co-translational insertion. The clear difference in EGFP fluorescence intensity seen in the avidin purified (proteo)polymersomes (figure 7.13) are clear proof of melittin’s ability to anchor soluble proteins to polymersomes. The fluorescence levels seen with melittin:EGFP before in the CF expression product, and after purification at nearly 40%, which match closely to the results seen with the avidin purification of the fluorescent polymersomes (figure 4.8b), confirm that significant amount of the expressed protein is indeed anchored on the polymersome membrane.

The work carried out has focused on examining the role of the melittin:EGFP fusion protein as an anchor for polymersome membrane insertion, and further work is required for the analysis of the other peptides; cecropin A and LAH4. Comparison of the rate of peptide insertion can be used to develop the ideal peptide sequence for polymersome anchoring through protein engineering techniques for further high-throughput screening and the targeted improvement of their binding capacities.

## Summary

Antimicrobial peptides have a key role in the development of polymersome membrane binding anchors. The simplicity of designing, cloning and expressing a wide range of proteins to these peptide sequences offers the potential for the surface decoration and functionalisation of polymersomes. These immobilised proteins have a wide range of applications in both bio-industrial processes and for environmental research offering enhanced protein display and stability.

Three antimicrobial peptides; cecropin A, melittin, and LAH4 were fused to EGFP, expressed in *E. coli* and successfully purified using affinity purification methods. Working with melittin:EGFP fusion protein, it was shown that SEC of the protein lead to two separation populations; one of aggregates, and another of soluble proteins. The presence of aggregates could be decreased by using detergents in the reconstitution protocol. Following the reconstitution of melittin:EGFP onto polymersome membranes, there is a shift in EGFP fluorescence intensities towards the polymersome separation range. Melittin:EGFP was shown also to reconstitute on the membranes of liposome GUVs during their electroformation. While melittin was also shown to assemble into pore forming complexes on liposomes and cause the leakage of vesicular contents. The use of fluorescence microscopy for the imaging of polymersomes is a challenge due to their much smaller sizes and alternative methods are needed for the examination of the effect of both melittin and melittin:EGFP fusion protein on polymersome membranes. Simpler and more direct route for the anchoring of soluble proteins on polymersome membranes was shown to be possible through the use of polymersome supplemented CF expression systems, and their subsequent purification.

Further work is needed to examine the binding capacities of melittin, and the other antimicrobial peptides, as well as the screening of engineered peptide sequences to develop the ideal peptide for surface decoration and functionalisation of polymersome membranes for a wide range of applications.

**Chapter 8**

# Discussion

8.1 INTRODUCTION

8.2 CLONING AND EXPRESSION OF CYTOCHROME P450 ENZYMES

8.3 PURIFICATION AND ACTIVITY OF CYP3A4 PROTEOPOLYMERSOMES

8.4 CHARACTERISATION OF PROTEOPOLYMERSOMES

8.5 PEPTIDE ANCHORS

8.6 SUMMARY

List of Abbreviations

ATP – Adenosine triphosphate

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

DSPE – 1,2-distearoyl-sn-glycero-3-phosphoethanolamine

EGFP – Enhanced green fluorescent protein

FAD – Flavin adenine dinucleotide

FMN – Flavin mononucleotide

PDMS – Polydimethylsiloxane

PEG – Polyethylene glycol

PEO – Polyethylene oxide

PLA – Polylactic acid

PMOXA – Poly-2-methyl-2-oxazoline

PTM – Post-translational modifications

SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

TEM – Transmission electron microscope

TMV – Tobacco mosaic virus

WT – Wild-type

List of Tables

**Table 8.1:** Protein expression systems utilised for the expression of CYP3A4, CPR and C*b5*.

List of Figures

**Figure 8.1:** Co-translational insertion of expressed membrane proteins onto polymersome membranes.

**Figure 8.2:** Chemical structures of biotin and desthiobiotin.

**Figure 8.3:** Cytochrome P450 monooxygenase enzymes crowding on polymersomes.

## Introduction

The work carried out in the previous chapters have helped elucidate and understand the role of polymersomes as a stabilising, biomimetic membrane environment for the cytochrome P450 microsomal monooxygenases. Starting with the aims of CF expression of the multi-enzyme system and its co-translational insertion into polymersomes, followed by its purification and characterisation, this project has explored new methods and tools to achieve its aims. Many of the methods and techniques used are different and contrary to most of the published work. The use of CF protein expression systems compared to the typically used bacterial and eukaryotic expression systems, the use of polymersomes instead of microsomes or liposomes highlight some of the novel ways research into CYP3A4 and microsomal monooxygenases has been carried out. These techniques have been used for the development of an *in vitro* multi-enzyme system that is purer and more stable than other established protocols, and with a much simpler and time-saving protocol. This preparation has opened avenues for the examination of the CYP3A4 monooxygenase and the elucidation of its structure, function and interaction and co-operativity with other microsomal monooxygenases and partner enzymes.

## Cloning and expression of cytochrome P450 enzymes

One of the main aims of this project was the co-translational insertion of the microsomal monooxygenases onto the polymersome membranes. To achieve this CF expression systems, supplemented with polymersomes were used for the expression of the enzymes. Typically, cell-based expression system systems are used for the expression of complex eukaryotic proteins such as CYP3A4, CPR and C*b5* (table 8.1). The CF expression of CYP3A4 and CPR has not been published before, yet with the many advances of CF expression systems in recent years, the use of these systems has become more and more advantageous. While *E. coli* CF systems typically struggle with eukaryotic protein expression (Endo and Sawasaki, 2006), as seen with the failed attempt at the expression of CYP3A4, the eukaryotic CF expression systems fare much better, as seen with the expression of CYP3A4 in figure 3.2.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 8.1:** Protein expression systems utilised for the expression of CYP3A4, CPR and C*b5*. | | | |
| **Enzyme** | **Expression system** | **Description** | **Ref.** |
| **CYP3A4 / CPR** | *E. coli* | Expression of WT and mutant variants, followed by purification and reconstitution. | (Gillam et al., 1993) |
| Human | Purification of liver cells microsomes | (Wang et al., 2000) |
| Insect | Coupled CYP3A4 and CPR expression in Sf9 insect cells using Baculovirus system, followed by purification of microsomes. | (Lee et al., 1995) |
| Yeast | Heterologous expression in *Saccharomyces cerevisiae* followed by purification of microsomes. | (Lamb et al., 2000) |
| **C*b5*** | *E. coli* | High level expression and purification of enzyme onto detergent. | (Holmans et al., 1994) |
| Wheat germ CF | Expression in liposome supplemented CF systems for co-translational insertion of protein. | (Nomura et al., 2008, Goren and Fox, 2008) |

While originally eukaryotic CF systems were utilised out of necessity and mostly only when expressing proteins of eukaryotic origin that were not readily expressed in the cheaper, simpler and higher yield *E. coli* systems, now many of these eukaryotic CF systems have their own advantages including processing protein PTMs, less protein aggregation and hosting the native environment of proteins. Wheat germ based CF systems similarly have many key advantages that allow it to compete directly with *E. coli* CF systems, including high yields, ability to do some PTMs and relatively low cost of production (Goshima et al., 2008, Takai and Endo, 2010). This high level of expression was indeed observed for the expression of CYP3A4 using the WEPRO wheat germ based CF system (figure 3.2c). For that reason, as well as lower costs and its ability to fold human proteins, the wheat germ CF system was used for further expression of CPR, C*b5*, and the combined multi-enzyme microsomal monooxygenase system.

The simple supplementation of the CF expression mixture, a key advantage offered by the open access CF expression systems, allows for the co-translational insertion of expressed proteins directly onto the polymersomes (figure 8.1), negating the need for any complex downstream reconstitution protocols. The openness of the CF system also allows for supplementation of key cofactors required for proper enzyme folding and functionality, including hemin for CYP3A4 and C*b5*, and FAD and FMN for CPR.

The WEPRO wheat germ CF system is really an ideal system for the expression of the cytochrome P450 enzymes. As well as successfully expressing CYP3A4 in high quantity, it has also successfully been used before for the expression of C*b5* (Goren and Fox, 2008). The WEPRO expression system is optimised for eukaryotic protein expression, with its two-phase expression system allowing for the high yield expression of proteins without the build-up of inhibitory elements and with the gradual exchange of the required translational elements *e.g.* amino acids, ATP, and cofactors, while the expression vector contains the E01 translational enhancer sequence. This non-natural untranslated sequence has shown to stimulate translation as much as the Ω sequence from TMV in wheat germ CF systems (Kamura et al., 2005, Ogawa et al., 2014).

Another key advantage of this system over the typical cell-based is the time savings. Expression can be carried from DNA to protein in a single day, and purified protein or proteopolymersome can be achieved in as little as two days, whereas with *E. coli* expression, the quickest of cell-based expression systems, it can take three days for the protein expression alone (considering transformation, overnight culturing and induction of expression).

|  |
| --- |
|  |
| **Figure 8.1:** Co-translational insertion of expressed membrane proteins onto polymersome membranes. **a)** RNA transcription stage, can be coupled to translation in same translation mixture or linked with transcription separate from translation. **b)** Protein translation stage, the peptide will interact with polymersome membranes present. **c)** Protein insertion, the hydrophobic membrane binding domain of the protein will insert into the polymersome membrane while the extravesicular portion will fold to form the active protein. |

With the publication of detailed and simple wheat germ CF extract preparation protocols (Takai and Endo, 2010), one of the major drawback to CF system utilisation, its cost, can be tackled. Preparation of lab made CF extract can be prepared at a fraction of the cost, and within only a day, while allowing for scale-up.

## Purification and activity of CYP3A4 proteopolymersomes

The wheat germ CF expressed CYP3A4 monooxygenase system is shown to be functional (figure 5.2 and 5.3). However, a lot of that activity is due to monooxygenase activity not present on the polymersome membranes, with activity levels in the absence of polymersomes reaching 70-80% of activity seen in the presence of polymersomes. While this shows that polymersomes, and especially their membranes, are required for enzyme binding and stabilisation and for functionality, there is also an alternative stabilising environment offered by the CF system. This stabilising environment of the CF system takes the enzymes away from the polymersomes, lowering the efficiency of protein insertion onto the polymersome membrane. With many of the enzyme present on other membranes this also lowers the overall activity seen when the proteopolymersomes are purified, and that is confirmed with the overall crude expression activity always being higher than that observed following any purification method. Yet with this high level of activity, this crude expression preparation is limited in its applications, filled with too many impurities from the CF expression system, as well as aggregated, misfolded and incomplete proteins lacking functionality.

*In vitro* enzyme preparations, such as the CYP3A4 proteopolymersome developed here, are required to have a high degree of purity to ensure accurate translation of functionality against its substrates and inhibitors as seen *in vivo*, as well as also giving an accurate picture of enzyme-enzyme interaction and structural analysis. Many impurities of the crude CF product, while unlikely, can possibly interact with the multi-enzyme CYP3A4 monooxygenase system and even inhibit its activity. Typically, with cell-based expression of CYP3A4 and other monooxygenase system involve a purification stage following expression and before reconstitution onto liposomes and other lipid based structures. For cytochrome enzymes expressed in *E. coli* purification techniques used include affinity (polyhistidine tag, and sepharose) and ion exchange chromatography (Farooq and Roberts, 2010). Similarly, when monooxygenase systems are expressed in eukaryotic cells on microsomes, they are purified using centrifugation methods to achieve high levels of purity (Lee et al., 1995, Wang et al., 2000). Purification from CF systems in principle are simpler and quicker, the openness of the system allow direct interaction with the proteins and the biomimetic membranes for direct purification without the requirement for cellular lysis or solubilisation. Since the enzymes insert into the polymersome membrane during the translation stage of expression, there is no need for a reconstitution step following purification.

Centrifugation, while highly applicable to lipid based membrane structures, is less suited for polymersomes. Liposomes have shown a propensity for pelleting during centrifugation, while it is resisted by the polymersomes (figures 4.1 and 4.3). Using that to our advantage and pelleting protein aggregates and CF extract impurities was an ineffective method. Most impurities remained after prolonged centrifugation as seen on the SDS-PAGE (figure 4.6, lanes 1 and 2), while the activity level following centrifugation (figures 5.6a and 5.7a) was around half of that observed in the crude CF expression product (figure 5.2 and 5.3). Similarly, filtration methods through the use of Vivaspin filtration columns do not provide pure enough samples, retaining many impurities of the CF extract as well as protein aggregates (figure 4.7), while also seemingly leading to a complete loss of any detectable activity (figure 5.8)

Alternative focus for purification shifted on to the proteins and polymersomes themselves rather than their environment. Protein purification while possible, the presence of polyhistidine tags on the enzymes allows for affinity purification, is not ideal. With the targeted purification of the folded proteins, it will be accompanied by the purification of protein aggregates and misfolded proteins. There is a risk also of purification of proteins off the polymersomes and causing the dissociation and unbinding of the proteins from the polymersome membrane. Targeting the polymersomes for purification is a much better strategy, with purification of only the polymersomes any protein not directly bound to the membrane is discarded. However, since polymersomes are composed of inert and unreactive polymers, the surface of polymersomes needs to be decorated in a way that allows the functionalisation of the polymersome surface and a target for the purification.

The decoration of polymersomes, with peptides, polysaccharides and lipids, has been carried out extensively (Egli et al., 2011b). The wide chemistry of polymers constituents of polymersomes allows for the conjugation of many chemical entities. The use of lipid (DSPE) conjugated PEG-biotin has been known to be possible. the hydrophobic environment of DSPE can stabilise itself within the hydrophobic region of diblock polymersomes. The extra-vesicular biotin molecule can act as a very effective affinity marker for avidin and neutravidin purification. The use of avidin resin has been shown to successfully purify biotinylated polymersomes, and when used to purify fluorescently labelled polymersomes, around 40% of polymersomes are purified and eluted, and at a higher concentration than at the start (figure 4.8). While 40% is a good rate of recovery, optimisation and improvement would be ideal. The use of a higher biotin concentration (from the current 5% molar ratio) would improve polymersome uptake (preventing polymersome loss of nearly 40%). However, since the biotin-avidin covalent bond is one of the strongest bonds, caution should be applied. The use of biotin alternatives e.g. desthiobiotin can offer a solution to this problem. With the same affinity to avidin yet with a weaker bond, elution can be simplified and optimised with this alternative. Desthiobiotin shares sufficient chemical structural similarity with biotin (figure 8.2) that it can be conjugated to DSPE-PEG or even more ideally to the diblock PBD-PEO through the use of carboxyl group binding to terminal amine group (EDC carbodiimide crosslinker can be used for activation of the amine group), and with the same polymer anchor for biotin as the composition of polymersomes uptake of the conjugate is more straight forward and efficient. Additionally, it can cut down on the elution time and biotin elution concentrations required significantly.

The avidin purification of the CYP3A4 monooxygenase proteopolymersome has produced a highly pure sample (Figure 5.9). There are no visible protein impurities, unlike those seen with the insect cell produced microsomal Baculosome (figure 5.11). However, the specific activity, a measure of a samples purity, is lower for the proteopolymersome than the Baculosome. The amount of enzyme in each sample has been quantified and it’s been shown that there is excess of each CYP3A4, CPR and C*b5* on the polymersomes. Yet however since the specific activity is lower there is less functional CYP3A4 enzyme. There are two main possible reasons for this; misfolded protein that lacks any activity, or since the monooxygenase system is a multi-enzyme system there is a lack of interaction between CYP3A4 and CPR or C*b5*, without which there is no complete electron transport chain. To have optimal levels of activity, there needs to be CPR directly interacting with each CYP3A4 enzyme. To improve the chances of this the “crowding” of enzymes on the polymersomes need to be improved (figure 8.3).

|  |  |
| --- | --- |
| **a)** | **b)** |
| **Figure 8.2:** Chemical structures of (**a**) biotin and (**b**) desthiobiotin. Biotin is conjugated to PEG through the terminal carboxyl group. | |

|  |
| --- |
|  |
| **Figure 8.3:** Cytochrome P450 monooxygenase enzymes crowding on polymersomes. **a)** High number of polymersomes and relatively low number of enzymes**. b)** Increasing the ratio of CPR to CYP3A4 improves the chance of interaction and giving activity. **c)** Decreasing the number of polymersomes while keeping the number of enzymes the same increases the chance for interactions. |

With enzymes too spread out on the polymersomes and no direct interaction, any activity is unlikely. There are two main methods to tackle this problem, the first is to increase the ratios of CPR and C*b5* to CYP3A4. When CPR is sufficiently in excess the crowding of the reductase is very high on all the vesicles so that any CYP3A4 enzyme available can interact with one (figure 8.3b). In the prepared CYP3A4 monooxygenase proteopolymersome the ratio of CYP3A4 to CPR is only 1:4, yet however in other CYP3A4 preparations, the optimal ratios have been shown to be up to 1:9 (Lee and Goldstein, 2012). There are limitations to this solution, the CF extract offers limited protein expression and any extra CPR expression will come at the expense of the expression of other enzymes, whether CYP3A4 or CPR. Alternatively, instead of increasing the ratio of CPR to CYP3A4, the number of polymersomes can be reduced. The reduction in number of polymersomes decreases the surface area available for protein insertion increasing the crowding of the enzymes. the downsides if this is with the decrease in polymersome count there is an increased chance of protein aggregating as it has no access to a membrane for binding and folding. The trade-off will be the increase in aggregation causing a decrease in the total activity of the proteopolymersomes, yet since only the (proteo)polymersomes are purified, the specific activity overall should be improved.

The absence of C*b5* also plays a key role in the low specific activity achieved. It’s been observed that C*b5* can boost CYP3A4 and other microsomal monooxygenase activity by several fold for multiple substrates (Porter, 2002). However, when tested with proteopolymersomes, the three-enzyme system which has C*b5* present, the total and specific activity is lower than the two-enzyme monooxygenase system without C*b5*. This can be explained as a result of the increase in both CYP3A4 and CPR expression, as well as the increase in the ratio of CPR to CYP3A4. It’s important to note that while C*b5* does boost CYP3A4 activity, it is only when present in close interaction with CPR as well, and that C*b5* interacting with CYP3A4 alone does not lead to activity as it lacks the required electron transfer to it, usually obtained from C*b5* reductase or CPR.

Another aspect to consider is the binding of monooxygenase system components. it has been shown that it is possible that microsomal monooxygenases such as CYP3A4 contain CPR binding sites (Williams et al., 2000) and with enzyme movement laterally across the membrane could drive the binding of CPR to CYP3A4 and increase the enzyme’s activity. These binding regions can be useful in the maintenance of the multi-enzyme complex in an ideally stable and functional conformation. With polymersomes, the polymer composition plays a key role in determining lateral movement (Discher and Ahmed, 2006). If such movement is desired then a suitable polymer must be chosen that can allow for it. The diblock polymer PBD-PEO, used for this research work, has a very tight arrangement. Alternative polymers however have been shown to have a looser arrangement, usually desired for their biodegradability and water permeability, polymers like PLA-PEO and PMOXA-PDMS-PMOXA can be modified to have the desired fluidity to allow for significant protein movement on its membrane. The length, and therefore the molecular weight, of the polymer have also been shown to affect the fluidity of the membrane. Having a more fluid and open membrane can also create a more receptive environment for protein insertion. Engineering of the proteopolymersome system such as the modification of polymer composition will always be required for optimal functionality.

Another way to engineer the system to allow for closer interactions between the enzyme components is through the engineering of the enzymes themselves. Although it is relatively still in its infancy stage, fusion cytochrome P450 enzymes and reductases have been expressed and shown to be active. Based on cytochrome P450 enzymes that natively have the reductase domain fused to the catalytic domain, such as CYP102A, microsomal monooxygenases have been fused to CPR including CYP3A4 (Gillam, 2007). However, with these fusion enzymes comes at a cost with decreased efficiency and overall activity compared to native systems.

A possible strategy to control the enzyme levels on the polymersomes is the reconstitution of enzymes on the polymersome membrane following CF expression in the presence of detergent. The detergent can be used as a stabiliser preventing enzyme aggregation. Polyhistidine tags are present on the enzymes and can be utilised for affinity purification, followed by dialysis in the presence of polymersomes for detergent removal and enzyme reconstitution onto the polymersome membrane. This strategy has been successfully employed with many of the cell expressed microsomal monooxygenases, and can be similarly utilised for CF expressed systems. Typically, CF expression is not utilised for purification and reconstitution protocols due to the lower expression yields observed, yet there is potential for utilising them for their time saving and easily modifiable environment. CF expression, purification and reconstitution can be done over a much shorter time scale than with most cell-based expression systems, and the presence of detergent throughout the expression stage ensures very limited aggregate formation and protein misfolding.

## Characterisation of proteopolymersomes

Polymersomes have two main advantages that make them more attractive compared to liposomes; their chemical diversity and their stability. While alternative biomimetics, such as liposomes and nanodiscs can offer similar function to polymersomes, they are inherently less stable. Since polymersomes are more stable they can remain a stabilising environment for the monooxygenases for much longer. This was observed when the activity of CYP3A4 on polymersomes was compared to that on the lipid Baculosome (figure 6.2). It was observed that after just a couple of weeks the activity on the Baculosome was almost lost, while it took almost twice as long for the same to happen on the proteopolymersomes. Yet since it is known that polymersomes can stably maintain their structure for weeks, the decrease in activity levels observed cannot be due to the vesicles, but rather due to the protein itself. Such decrease in activity levels and protein degradation is usually caused by proteases, and while protease inhibitors can utilised to minimise their effect they also have an undesired side-effect of inhibiting CYP3A4 monooxygenase activity. The use of antimicrobials to prevent the growth of any microbes, the main producers of the proteases, however can minimise any protein degradation and loss of activity (figure 6.4). Proteopolymersomes can maintain their activity levels for weeks when microbial growth is inhibited and even show an increase activity (10 – 20%) over the first week. This increase in activity levels is an interesting observation, possibly hinting that it can take time for the enzymes to reach each other on the membrane surface and interact due to the low lateral fluidity of the polymer membrane, as discussed previously. It would be interesting to observe whether changing the polymer composition will influence the delay of the spike in activity, with looser and fluid membranes seeing the peak in activity sooner.

The flow cytometry analysis offers great insight into the proteopolymersome system. Typically, most flow cytometry analysis of polymersomes has revolved around their use as carrier vesicles (Nallani et al., 2009, Katz et al., 2010), These strategies utilise polymersomes packed with fluorescent molecules making them easy to track. To track the membrane bound proteins though, the enzymes were labelled using fluorescent antibodies. Similarly, the polymersomes themselves were labelled with fluorescent rhodamine conjugated to lipid. This labelling of polymersomes has highlighted the efficiency on lipid incorporation onto the polymersome membrane. Around 40% of polymersomes show little to no fluorescence (Figure 6.6). This shows that while polymersomes are capable of lipid uptake, it is not perhaps not as efficient enough that a majority of vesicles would take up the lipids. This is important to consider since the polymersomes were themselves purified using lipid-conjugated biotin tags, and with similar incorporation efficiencies, 40% of proteopolymersomes are lost and not purified simply due to being untagged. The use of biotin conjugated directly to polymers instead of to lipid can offer a better alternative, and it has been successfully carried out with biotin conjugated to PMOXA polymer in a triblock (PMOXA-PDMS-PMOXA) arrangement (Broz et al., 2005).

The fluorescent protein tells an interesting story too, with the discovery that the majority of polymersomes do not contain incorporated proteins on their membrane (figure 6.9), or at least at an undetectable level. In an multi-enzyme system, multiple enzymes are required on a single polymersome for functionality, and having nearly 70% of polymersomes with no or few enzymes is an inhibitor to activity. The specific activity calculated is lower than it is expected to be, and the lack of crowding of the enzymes on the polymersomes was discussed as the main cause of that, and seeing the comparison of fluorescence between empty polymersomes and the (proteo)polymersomes it would lend weight behind that theory. A similar story is told by the TEM imaging of the (proteo)polymersomes. We see that while there is immunogold labelling of the enzymes on the membrane of polymersomes, there are many polymersomes that remain unlabelled. Ideally you would see high level labelling of all polymersomes indicating the presence of proteins on their membranes.

## Peptide anchors

Polymersomes have always been attractive because of their ease of modification and alteration. A lot of work has been published on the functionalisation of polymersomes through decoration of the membrane surface through the conjugation of ligands to either polymers or fully formed polymersome vesicles, as well as the use of polymers with desirable functional group for vesicle formation (Egli et al., 2011b). Another underexplored avenue for functionalization of polymersomes is the use of anchoring peptide sequences that can decorate the surface of polymersome with any desired amino acid sequence. Antimicrobial peptides are a very attractive option for these peptide anchors as they specifically target the hydrophobic environment of membranes. These peptides, along with other membrane spanning regions of proteins have been shown to successfully reconstitute onto the polymersome membrane (Vijayan et al., 2005). The melittin antimicrobial peptide is an ideal example of a such a polymersome membrane anchor. Reconstitution and co-translational insertion of the peptide was possible (figures 7.7 and 7.13), even when the peptide was fused to large soluble, non-membrane binding protein, such as EGFP. By comparison liposomes have shown that they can as well attract antimicrobial peptides, however the stability of the vesicles are put in jeopardy. Melittin, and other antimicrobial peptides, have been shown to destabilise and lyse liposomes (Vijayan et al., 2005), as well as cause the leakage of their contents (figures 7.10 and 7.11). The challenge from now is the further examination in detail of these antimicrobial peptide and the analysis of their applications in terms of polymersome decoration with soluble and typically non-membrane bound proteins, as well as anchors for membrane protein that face challenges in membrane binding.

## Summary

The study of membrane proteins has always been a difficult challenge due to their hydrophobicity and requirement for an equally hydrophobic and stabilising environment. CYP3A4, like other cytochrome P450 enzymes, is no different in its requirements, and as such most research on the enzyme originates from complicated, time consuming and often unstable lipid based membranes. The use of CF for co-translational insertion of the multi-enzyme CYP3A4 monooxygenase system has been successfully carried out to create functional CYP3A4 monooxygenase proteopolymersomes at a fraction of the time. Modification of the polymersome membrane has allowed for the simple purification of the proteopolymersome to a very high degree, coupled with the use of supplements for the creation of proteopolymersome preparations with much superior levels of stability than those seen with lipid vesicles. New polymersome modification methods has also been developed using membrane binding peptides that can stably anchor larger soluble proteins and decorate the surface of polymersomes, showing advantages of polymersomes over liposomes in its ability to be modified and altered according to need.

**Chapter 9**

# Future Perspectives and Concluding Remarks

9.1 INTRODUCTION

9.2 IMPROVING THE CYP3A4 PROTEOPOLYMERSOME

9.3 UTILISING THE CYP3A4 PROTEOPOLYMERSOME

9.4 CONCLUSION

List of Abbreviations

C*b5* – Cytochrome *b5*

CF – Cell free

CPR – Cytochrome P450 reductase

CYP3A4 – Cytochrome P450 3A4

GPCR – G protein coupled receptor

MMO – Microsomal monooxygenase

NMR – Nuclear magnetic resonance

PBD – Polybutadiene

PDMS – Polydimethylsiloxane

PEO – Polyethylene oxide

PMOXA – Poly-2-methyl-2-oxazoline

uAA – Unnatural amino acid

## Introduction

The exploration of the role that polymersomes play in offering biomimetic environments for membrane binding proteins is still in its early stages. However, many significant and highly promising advances have been achieved. Complex and difficult membrane proteins with multiple membrane spanning domains such as GPCRs have been successfully reconstituted and co-translationally inserted onto polymersome membranes (Nallani et al., 2011, May et al., 2013). Polymersomes themselves have also been advanced, the standard utilised polymers such as PBD-PEO and PMOXA-PDMS-PMOXA have been altered and modified for specific application, as well as the mixing of both lipids and polymers for the creation of hybrid vesicles with ideal features from both. There is still a lot that needs to be explored in the polymersome up take of proteins, especially within CF expression systems, optimising protein yield, binding and folding to create highly functional systems. There is also room for improving the characterisation and purification yield of the system and scaling it upwards.

## Improving the CYP3A4 Proteopolymersome

One of the main goals for improving the CYP3A4 monooxygenase proteopolymersome is the increase of protein yield. The increase in protein yields will lead to higher levels of total activity. The higher the yield of proteins, the more crowded the proteopolymersomes are, aiding greatly in increasing the specific activity of the proteopolymersomes. With CF expression systems, it is a given that yields will be low, especially compared to cell-based expression systems. However, some attempts can be carried out to optimise yields. One of the main ways to carry this out is to produce CF expression system within a lab setting. Protocols for the higher yield wheat germ CF system has already been published (Takai and Endo, 2010), and while only requiring basic chemicals and lab equipment. The successful utilisation of such lab-made expression systems can cut down the cost and time requirements needed with commercial CF expression systems. An important factor with the decrease in running costs is the ability for scale-up. Without it being a limitation, scale-up of the expression system can be used to increase yields significantly with minimal increase in costs.

Another method for the optimisation of yields could be achieved through the reconstitution of CF expressed and purified proteins. Currently in the system used, only a fraction of expressed proteins inserts into the polymersomes, and a fraction of those proteopolymersomes are then purified to give the pure proteopolymersomes preparation that is utilised. To minimise the loss of proteins and maximise their presence on the polymersomes an alternative approach to co-translational insertion can be taken. While still utilising CF expression systems, the expression mixture can be supplemented with detergent instead of polymersomes which can act as the stabilising environment for the protein and hold them in folded and non-aggregated conformation. Subsequently, the purification of these detergent-protein complexes, *e.g.* through the use of polyhistidine tag affinity purification, can purify the protein to a much higher degree of efficiency than that obtained with the current avidin-biotin purification. The reconstitution of the purified proteins can be carried out as have been done previously with CYP3A4 (Gillam et al., 1993, Grinkova et al., 2010), and with polymersomes (Meier et al., 2000, Kumar et al., 2007, Beales et al., 2017). The removal of detergent from the detergent-protein complex in presence of polymersomes, *e.g.* using dialysis or bio-beads, will drive the insertion of the folded proteins off the detergent onto the polymersome membrane.

As discussed earlier there is a high level of variation among the polymer constituents of polymersomes, and that variety on offer can itself be taken advantage of to produce the ideal polymersomes for the desired application. Currently, the polymer used is the compact and relatively non-fluid PBD-PEO. The entanglement of the polymer as a membrane limits its fluidity and lateral mobility (Discher and Ahmed, 2006), a detriment to multi-enzyme system such as that of microsomal monooxygenases which require movement towards each other, interactions and complete electron chain formation for functionality. The length of polymer, and as a result its molecular weight, plays a role in polymer entanglement and membrane fluidity and lateral diffusion capability. PMOXA-PDMS based polymers have shown a decrease in lateral diffusion with increasing polymer chain length (Itel et al., 2014). It is important to note that fluid membranes are not only a concern for the lateral diffusion of proteins across the membrane to enable protein-protein interactions, but also as an aid in insertion. Looser arrangement of polymers can provide a more accepting environment for membrane proteins inserting themselves into membranes. Screening of various polymers with varying molecular weights can be done to observe the yields of purified proteins on proteopolymersomes (to identify protein insertion capacities), as well as observing activity (to identify the role of membrane fluidity and protein diffusion). Finally, one key alteration to polymers is its conjugation to functional groups, such as biotin or desthiobiotin, that can aid in its purification. The conjugation of biotin to polymers can increase the uptake of the functional group into the polymersome membranes during polymer extrusion, optimising the purification of the vesicles by the avidin resin. While the conjugation of desthiobiotin can offer an additional advantage by optimising the elution off the avidin resin.

## Utilising the CYP3A4 Proteopolymersome

While the CYP3A4 monooxygenase proteopolymersome still has room for improvement, it is within its own right a significant achievement with its own benefits and applications. The first of which is the study of the CYP3A4 itself as a functional enzyme unit. The most attractive feature of CYP3A4 is its broad substrate specificity which it translates into metabolism against a wide range of endogenous, as well as exogenous drug substrates. Most of the work on understanding this broad substrate promiscuity has focused on the F-G helices region (F, F’, G, G’ and the connecting loops) of CYP3A4 which forms the upper wall of the active site cavity (Ekroos and Sjogren, 2006, Gay et al., 2010). The flexibility and plasticity of these helices allows the modification of the enzyme conformation, and the acceptance of a range of ligands requiring larger binding cavities such as ketoconazole and erythromycin, especially by the F and F’ helices. Additionally, these helices can decrease the substrate cavity causing the entrapment and immobilisation of the ligands (Sevrioukova and Poulos, 2017). The movement of these helices trigger modifications of other surrounding helices (B’, H and I) that alter both enzyme conformation and activity. The F-G region plays a key role in ligand binding in many other cytochrome P450s, including 2C9, 2B6, 2A6 and 3A7 (Torimoto et al., 2007, Gay et al., 2010). While the structures of CYP3A4, and other cytochrome P450s are being elucidated at a growing rate, there is still difficulty faced in the use of NMR techniques and X-ray crystallography of low solubilisation and stability (Gay et al., 2010). The use of cytochrome P450 proteopolymersomes can offer a solution for the study of these important enzymes, especially in enzyme-ligand complexes that will aid in understanding the metabolism of pharmaceutical drug targets and the design of ideal targets.

Like the helices, polymorphisms in the CYP3A4 sequence can also alter the ligand specificity. Genetic polymorphism prevalence varies greatly across populations and the study of both the prevalence of a polymorphism as well as its effect on ligand binding and metabolism is vital for understanding and designing novel ligand for targeted and specialised ligand for personalised medicine (Lee et al., 2013a, Klein and Zanger, 2013). The use of proteopolymersomes can offer a method for fast, high-throughput screening of polymorphisms within CYP3A4. The use of CYP3A4 proteopolymersomes for structural and metabolic studies is a great aid in the development of designed, personalised, targeted and functional drug molecules and open new pathways for drug development and understanding drug side-effects.

Proteopolymersomes also offer an attractive platform for the use of protein engineering techniques. altering the structure of the enzyme can be applied advantageously to improve expression, stability, activity, or indeed introduce new functionality. Coupled with CF technology, proteopolymersomes offer a unique new platform for protein engineering allowing for the manipulation of protein sequence, uAAs incorporation and direct assaying using toxic ligands. CF expression systems, such as that of the wheat-germ allow for high-throughput screening of enzymes (Kanno and Tozawa, 2010). The open-access to the expression system can allow coupling of genotype to the phenotype for high throughput screening of engineered enzymes providing new and desirable characteristics for biotechnological applications.

Another possible desirable aspect of CYP3A4 open for engineering is its interaction with CPR. Alteration of both the enzyme-enzyme interactions, for tighter and more efficient binding and interactions, or the fusion of the reductase domain to CYP3A4 are both very advantageous. The main feature of the CYP3A4 proteopolymersome is simply its ability to be used for the study of all the relevant ligand metabolism. This proteopolymersome is a useful tool for the elucidation of CYP3A4 substrates and inhibitors, as well as a tool for the novel discovery of previously unknown ligands and possibly the discovery of the role of CYP3A4 in new metabolic systems.

Finally, and quite importantly, the CYP3A4 proteopolymersome as an *in vitro* enzyme platform is applicable to all the other members of the MMO enzyme family. These enzymes share key structural features (*e.g.* N-terminus membrane spanning domain and large cytoplasmic c-terminus domain) that allow for similar proteopolymersome development for other MMOs. Additionally, it offers a platform for the study of CYP3A4 protein-protein interactions, not only with their reductases, but also to other members of MMOs. It is known that MMOs exist in multi-enzyme complexes, both as homomeric complexes of the same enzyme, and as a heteromeric complexes of multiple different enzymes. The study of these interactions and how they work will help understand how physiological metabolism of many important ligands and drugs occurs and what effect this co-operativity between the same and different enzymes have, and can answer many questions for the development of novel drug targets.

## Conclusion

Polymersomes are membrane protein solubilising and stabilising platform with exciting potential. Their large variability and stability offer advantages not possible with other membrane protein environments. The use of CF expression systems for the co-translational insertion of multi-enzyme system onto polymersomes offers a unique, simple and quick method for the preparation of functional in vitro enzyme preparations. Aided by the ability to introduce variability within the membranes of polymersomes, their purification is also simple and quicker than standard protocols. The produced CYP3A4 proteopolymersome was not only proven to be functional, but also highly pure. The characterisation of the proteopolymersome yielded promising data confirming protein insertion and stability. This proteopolymersome is a valuable tool in the study of not only CYP3A4, but other MMOs, for understanding metabolic processes, drug design, development of novel targets and side effects.

# Publications

“Development of the CYP3A4 microsomal monooxygenase proteopolymersome”. *Article in preparation.*

**Omar Ali, H.**, Alessa, A. H. A., Al-nuaemi, I. J., Wong, T. S. “Protein engineering for lignocellulose degradation.” Environmental Science and Engineering Volume 8 (2017). 1st ed. USA: Studium Press LLC.

Warburton, M., **Omar Ali, H.**, Wai Choon, L., Othusitse A. M., Abdullah Zubir A. Z., Maddock, S., and Wong, T. S. "OneClick: a program for designing focused mutagenesis experiments." *AIMS Bioengineering* 2 (2015): 126-143.

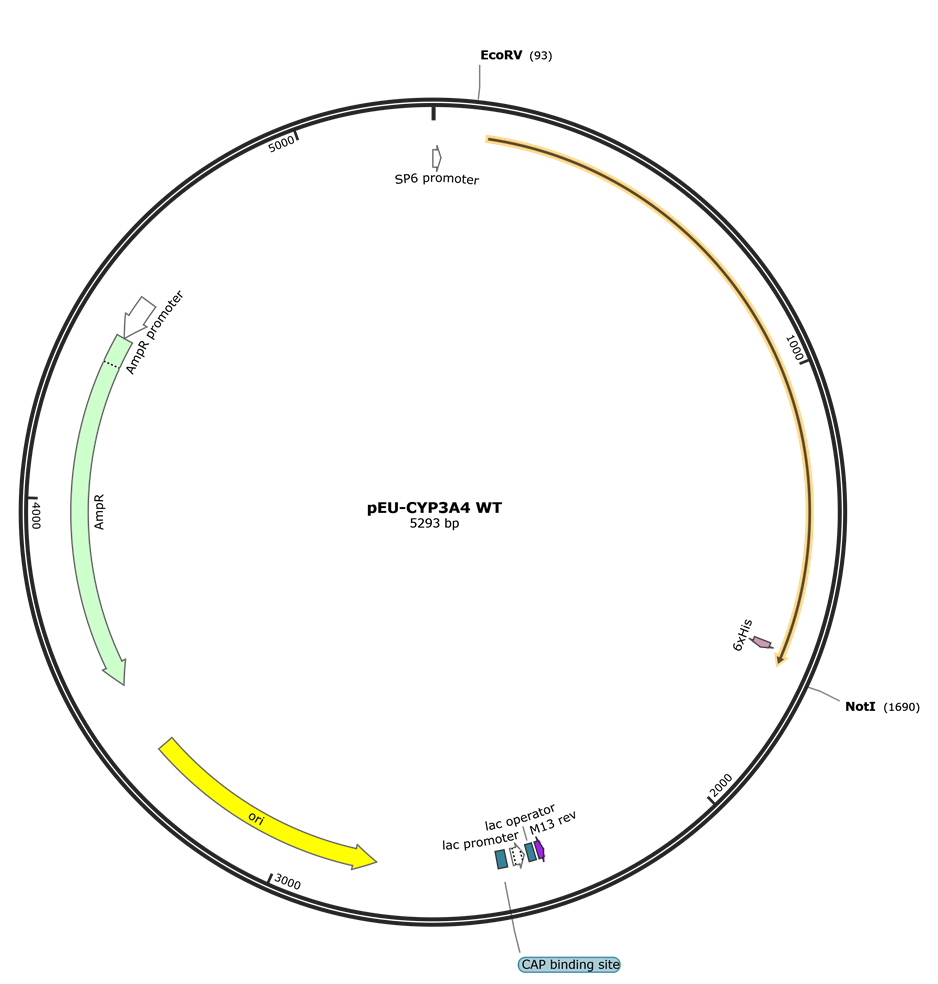
Jajesniak, P., **Omar Ali, H.**, and Wong, T. S. "Carbon dioxide capture and utilization using biological systems: opportunities and challenges." *Journal of Bioprocessing & Biotechniques* 4, no. 3 (2014): 1.

# Appendix

1. CYP3A4 DNA sequence (highlighted sequence absent in Nf14 construct)

ATGGCT**CTGATTCCGGATCTGGCGATGGAAACCTGG**CTGTTATTAGCAGTTTCCCTGGTGCTCCTCTATCTATATGGAACCCATTCACATGGACTTTTTAAGAAGCTTGGCATTCCAGGGCCCACACCTCTGCCTTTTTTGGGAAATATTTTGTCCTACCATAAGGGCTTTTGTATGTTTGACATGGAATGTCATAAAAAGTATGGAAAAGTGTGGGGCTTTTATGATGGTCAACAGCCTGTGCTGGCTATCACAGATCCTGACATGATCAAAACAGTGCTAGTGAAAGAATGTTATTCTGTCTTCACAAACCGGAGGCCTTTTGGTCCAGTGGGATTTATGAAAAGTGCCATCTCTATAGCTGAGGATGAAGAATGGAAGAGATTACGATCATTGCTGTCTCCAACCTTCACCAGTGGAAAACTCAAGGAGATGGTCCCTATCATTGCCCAGTATGGAGATGTGTTGGTGAGAAATCTGAGGCGGGAAGCAGAGACAGGCAAGCCTGTCACCTTGAAAGACGTCTTTGGGGCCTACAGCATGGATGTGATCACTAGCACATCATTTGGAGTGAACATCGACTCTCTCAACAATCCACAAGACCCCTTTGTGGAAAACACCAAGAAGCTTTTAAGATTTGATTTTTTGGATCCATTCTTTCTCTCAATAACAGTCTTTCCATTCCTCATCCCAATTCTTGAAGTATTAAATATCTGTGTGTTTCCAAGAGAAGTTACAAATTTTTTAAGAAAATCTGTAAAAAGGATGAAAGAAAGTCGCCTCGAAGATACACAAAAGCACCGAGTGGATTTCCTTCAGCTGATGATTGACTCTCAGAACTCAAAAGAAACTGAGTCCCACAAAGCTCTGTCCGATCTGGAGCTCGTGGCCCAATCAATTATCTTTATTTTTGCTGGCTATGAAACCACGAGCAGTGTTCTCTCCTTCATTATGTATGAACTGGCCACTCACCCTGATGTCCAGCAGAAACTGCAGGAGGAAATTGATGCAGTTTTACCCAATAAGGCACCACCCACCTATGATACTGTGCTACAGATGGAGTATCTTGACATGGTGGTGAATGAAACGCTCAGATTATTCCCAATTGCTATGAGACTTGAGAGGGTCTGCAAAAAAGATGTTGAGATCAATGGGATGTTCATTCCCAAAGGGGTGGTGGTGATGATTCCAAGCTATGCTCTTCACCGTGACCCAAAGTACTGGACAGAGCCTGAGAAGTTCCTCCCTGAAAGATTCAGCAAGAAGAACAAGGACAACATAGATCCTTACATATACACACCCTTTGGAAGTGGACCCAGAAACTGCATTGGCATGAGGTTTGCTCTCATGAACATGAAACTTGCTCTAATCAGAGTCCTTCAGAACTTCTCCTTCAAACCTTGTAAAGAAACACAGATCCCCCTGAAATTAAGCTTAGGAGGACTTCTTCAACCAGAAAAACCCGTTGTTCTAAAGGTTGAGTCAAGGGATGGCACCGTAAGTGGAGCGTCGACCCATCATCATCATCATCATTAA

B) pEU-CYP3A4 plasmid map

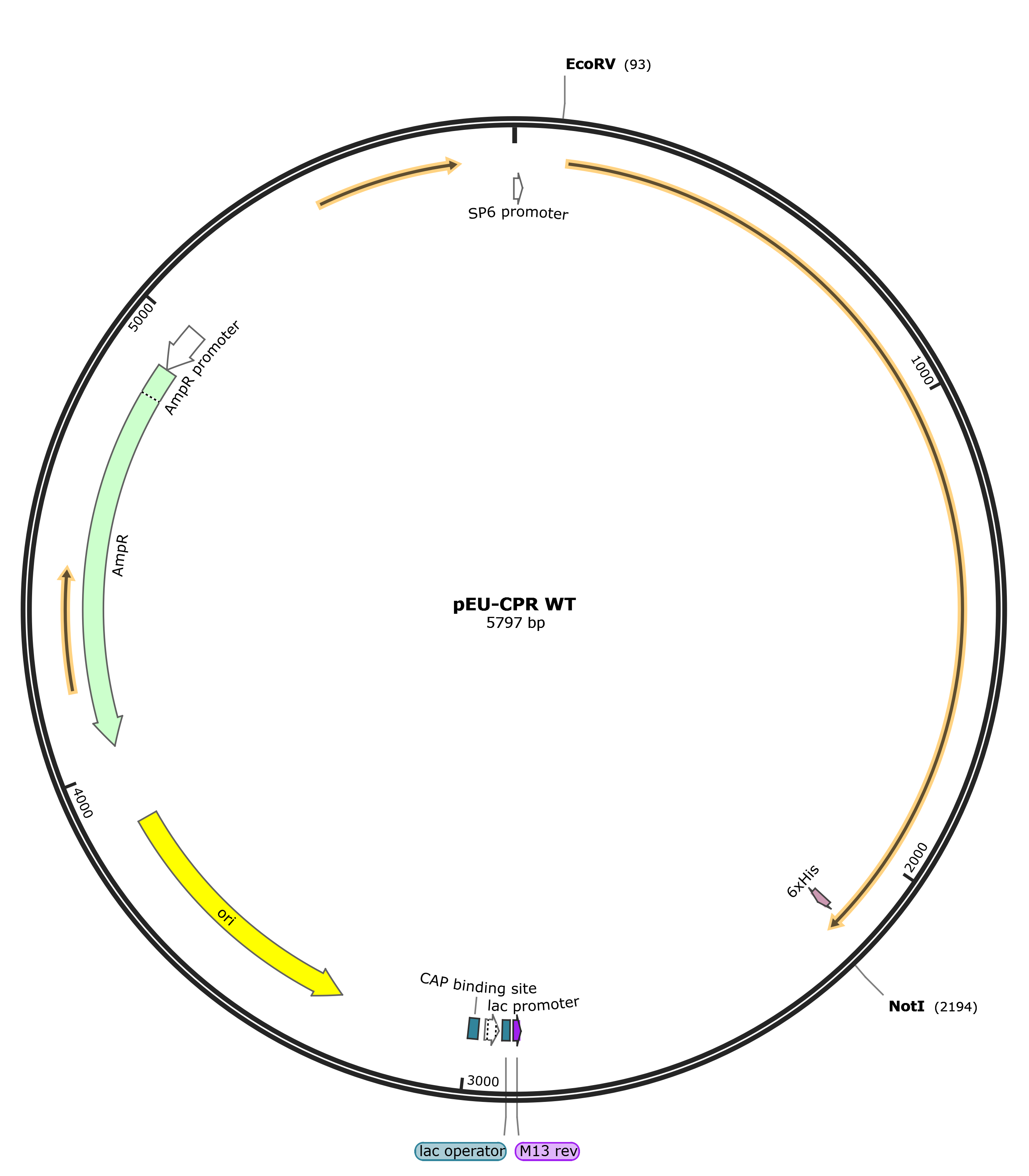


**CYP3A4**

1. CPR DNA sequence (with highlighted C-terminus 1D4 tag)

ATGGGCGACAGCCACGTGGACACCAGCAGCACCGTGAGCGAGGCCGTGGCCGAGGAGGTGAGCCTGTTCAGCATGACCGACATGATCCTGTTCAGCCTGATCGTGGGCCTGCTGACCTACTGGTTCCTGTTCAGGAAGAAGAAGGAGGAGGTGCCCGAGTTCACCAAGATCCAGACCCTGACCAGCAGCGTGAGGGAGAGCAGCTTCGTGGAGAAGATGAAGAAGACCGGCAGGAACATCATCGTGTTCTACGGCAGCCAGACCGGCACCGCCGAGGAGTTCGCCAACAGGCTGAGCAAGGACGCCCACAGGTACGGCATGAGGGGCATGAGCGCCGACCCCGAGGAGTACGACCTGGCCGACCTGAGCAGCCTGCCCGAGATCGACAACGCCCTGGTGGTGTTCTGCATGGCCACCTACGGCGAGGGCGACCCCACCGACAACGCCCAGGACTTCTACGACTGGCTGCAGGAGACCGACGTGGACCTGAGCGGCGTGAAGTTCGCCGTGTTCGGCCTGGGCAACAAGACCTACGAGCACTTCAACGCCATGGGCAAGTACGTGGACAAGAGGCTGGAGCAGCTGGGCGCCCAGAGGATCTTCGAGCTGGGCCTGGGCGACGACGACGGCAACCTGGAGGAGGACTTCATCACCTGGAGGGAGCAGTTCTGGCCCGCCGTGTGCGAGCACTTCGGCGTGGAGGCCACCGGCGAGGAGAGCAGCATCAGGCAGTACGAGCTGGTGGTGCACACCGACATCGACGCCGCCAAGGTGTACATGGGCGAGATGGGCAGGCTGAAGAGCTACGAGAACCAGAAGCCCCCCTTCGACGCCAAGAACCCCTTCCTGGCCGCCGTGACCACCAACAGGAAGCTGAACCAGGGCACCGAGAGGCACCTGATGCACCTGGAGCTGGACATCAGCGACAGCAAGATCAGGTACGAGAGCGGCGACCACGTGGCCGTGTACCCCGCCAACGACAGCGCCCTGGTGAACCAGCTGGGCAAGATCCTGGGCGCCGACCTGGACGTGGTGATGAGCCTGAACAACCTGGACGAGGAGAGCAACAAGAAGCACCCCTTCCCCTGCCCCACCAGCTACAGGACCGCCCTGACCTACTACCTGGACATCACCAACCCCCCCAGGACCAACGTGCTGTACGAGCTGGCCCAGTACGCCAGCGAGCCCAGCGAGCAGGAGCTGCTGAGGAAGATGGCCAGCAGCAGCGGCGAGGGCAAGGAGCTGTACCTGAGCTGGGTGGTGGAGGCCAGGAGGCACATCCTGGCCATCCTGCAGGACTGCCCCAGCCTGAGGCCCCCCATCGACCACCTGTGCGAGCTGCTGCCCAGGCTGCAGGCCAGGTACTACAGCATCGCCAGCAGCAGCAAGGTGCACCCCAACAGCGTGCACATCTGCGCCGTGGTGGTGGAGTACGAGACCAAGGCCGGCAGGATCAACAAGGGCGTGGCCACCAACTGGCTGAGGGCCAAGGAGCCCGCCGGCGAGAACGGCGGCAGGGCCCTGGTGCCCATGTTCGTGAGGAAGAGCCAGTTCAGGCTGCCCTTCAAGGCCACCACCCCCGTGATCATGGTGGGCCCCGGCACCGGCGTGGCCCCCTTCATCGGCTTCATCCAGGAGAGGGCCTGGCTGAGGCAGCAGGGCAAGGAGGTGGGCGAGACCCTGCTGTACTACGGCTGCAGGAGGAGCGACGAGGACTACCTGTACAGGGAGGAGCTGGCCCAGTTCCACAGGGACGGCGCCCTGACCCAGCTGAACGTGGCCTTCAGCAGGGAGCAGAGCCACAAGGTGTACGTGCAGCACCTGCTGAAGCAGGACAGGGAGCACCTGTGGAAGCTGATCGAGGGCGGCGCCCACATCTACGTGTGCGGCGACGCCAGGAACATGGCCAGGGACGTGCAGAACACCTTCTACGACATCGTGGCCGAGCTGGGCGCCATGGAGCACGCCCAGGCCGTGGACTACATCAAGAAGCTGATGACCAAGGGCAGGTACAGCCTGGACGTGTGGAGC**ACCGAGACCTCCCAGGTGGCCCCCGCC**TAA

1. pEU-CPR plasmid map

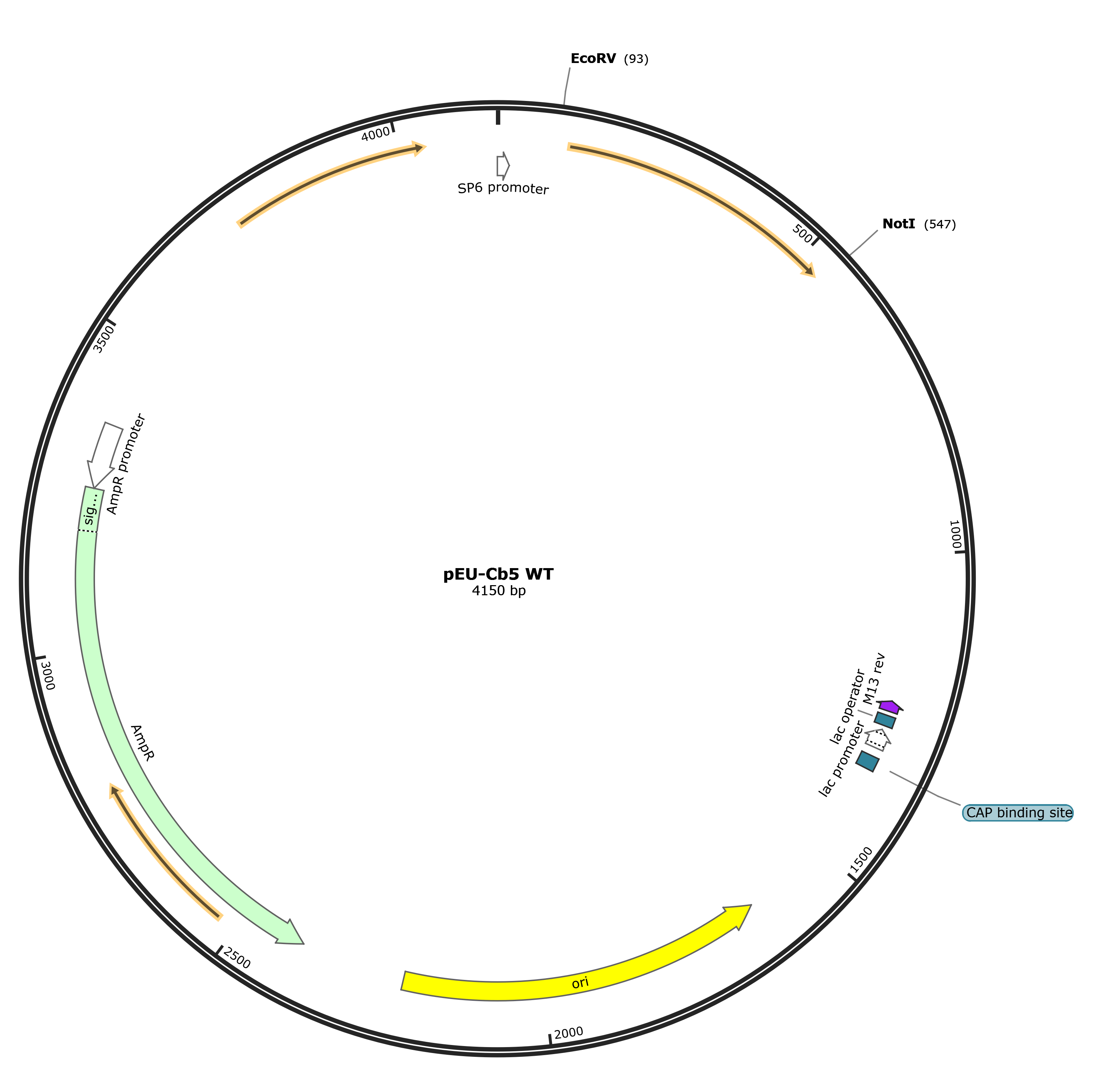


**CPR**

1. Cb5 DNA sequence (with highlighted C-terminus 1D4 tag)

ATGGCCGAGCAGAGCGACGAGGCCGTGAAGTACTACACCCTGGAGGAGATCCAGAAGCACAACCACAGCAAGAGCACCTGGCTGATCCTGCACCACAAGGTGTACGACCTGACCAAGTTCCTGGAGGAGCACCCCGGCGGCGAGGAGGTGCTGAGGGAGCAGGCCGGCGGCGACGCCACCGAGAACTTCGAGGACGTGGGCCACAGCACCGACGCCAGGGAGATGAGCAAGACCTTCATCATCGGCGAGCTGCACCCCGACGACAGGCCCAAGCTGAACAAGCCCCCCGAGACCCTGATCACCACCATCGACAGCAGCAGCAGCTGGTGGACCAACTGGGTGATCCCCGCCATCAGCGCCGTGGCCGTGGCCCTGATGTACAGGCTGTACATGGCCGAGGAC**ACCGAGACCTCCCAGGTGGCCCCCGCC**TAA

1. ***pEU-***Cb5 plasmid map



**C*b5***

1. BOMCC product fluorescent standard curve (with trendline)

|  |
| --- |
|  |

# References

AHN, J. H., CHU, H. S., KIM, T. W., OH, I. S., CHOI, C. Y., HAHN, G. H., PARK, C. G. & KIM, D. M. 2005. Cell-free synthesis of recombinant proteins from PCR-amplified genes at a comparable productivity to that of plasmid-based reactions. *Biochem Biophys Res Commun,* 338**,** 1346-52.

AHN, J. H., KANG, T. J. & KIM, D. M. 2008. Tuning the expression level of recombinant proteins by modulating mRNA stability in a cell-free protein synthesis system. *Biotechnol Bioeng,* 101**,** 422-7.

ANDREASSON-OCHSNER, M., FU, Z. K., MAY, S., XIU, L. Y., NALLANI, M. & SINNER, E. K. 2012. Selective Deposition and Self-Assembly of Triblock Copolymers into Matrix Arrays for Membrane Protein Production. *Langmuir,* 28**,** 2044-2048.

ARIFIN, D. R. & PALMER, A. F. 2005. Polymersome encapsulated hemoglobin: a novel type of oxygen carrier. *Biomacromolecules,* 6**,** 2172-81.

AVENAUD, P., CASTROVIEJO, M., CLARET, S., ROSENBAUM, J., MEGRAUD, F. & MENARD, A. 2004. Expression and activity of the cytolethal distending toxin of Helicobacter hepaticus. *Biochem Biophys Res Commun,* 318**,** 739-45.

BARNES, H. J., ARLOTTO, M. P. & WATERMAN, M. R. 1991. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in Escherichia coli. *Proc Natl Acad Sci U S A,* 88**,** 5597-601.

BATYCKY, R. P., HANES, J., LANGER, R. & EDWARDS, D. A. 1997. A theoretical model of erosion and macromolecular drug release from biodegrading microspheres. *J Pharm Sci,* 86**,** 1464-77.

BAYBURT, T. H., CARLSON, J. W. & SLIGAR, S. G. 1998. Reconstitution and imaging of a membrane protein in a nanometer-size phospholipid bilayer. *J Struct Biol,* 123**,** 37-44.

BEALES, P. A., KHAN, S., MUENCH, S. P. & JEUKEN, L. J. 2017. Durable vesicles for reconstitution of membrane proteins in biotechnology. *Biochem Soc Trans,* 45**,** 15-26.

BECHINGER, B. 1996. Towards membrane protein design: pH-sensitive topology of histidine-containing polypeptides. *J Mol Biol,* 263**,** 768-75.

BECHINGER, B. 1997. Structure and functions of channel-forming peptides: magainins, cecropins, melittin and alamethicin. *J Membr Biol,* 156**,** 197-211.

BERKMEN, M. 2012. Production of disulfide-bonded proteins in Escherichia coli. *Protein Expr Purif,* 82**,** 240-51.

BERMUDEZ, H., BRANNAN, A. K., HAMMER, D. A., BATES, F. S. & DISCHER, D. E. 2001. Molecular weight dependence of polymersome membrane elasticity and stability. *arXiv preprint cond-mat/0110088*.

BERRIER, C., GUILVOUT, I., BAYAN, N., PARK, K. H., MESNEAU, A., CHAMI, M., PUGSLEY, A. P. & GHAZI, A. 2011. Coupled cell-free synthesis and lipid vesicle insertion of a functional oligomeric channel MscL MscL does not need the insertase YidC for insertion in vitro. *Biochim Biophys Acta,* 1808**,** 41-6.

BISHOP-BAILEY, D., THOMSON, S., ASKARI, A., FAULKNER, A. & WHEELER-JONES, C. 2014. Lipid-metabolizing CYPs in the regulation and dysregulation of metabolism. *Annual review of nutrition,* 34**,** 261-279.

BOYER, M. E., STAPLETON, J. A., KUCHENREUTHER, J. M., WANG, C. W. & SWARTZ, J. R. 2008. Cell-free synthesis and maturation of [FeFe] hydrogenases. *Biotechnol Bioeng,* 99**,** 59-67.

BRODEL, A. K., SONNABEND, A. & KUBICK, S. 2014. Cell-free protein expression based on extracts from CHO cells. *Biotechnol Bioeng,* 111**,** 25-36.

BRODEL, A. K., SONNABEND, A., ROBERTS, L. O., STECH, M., WUSTENHAGEN, D. A. & KUBICK, S. 2013. IRES-mediated translation of membrane proteins and glycoproteins in eukaryotic cell-free systems. *PLoS One,* 8**,** e82234.

BROZ, P., BENITO, S. M., SAW, C., BURGER, P., HEIDER, H., PFISTERER, M., MARSCH, S., MEIER, W. & HUNZIKER, P. 2005. Cell targeting by a generic receptor-targeted polymer nanocontainer platform. *J Control Release,* 102**,** 475-88.

BUI, H. T., URNAKOSHI, H., NGO, K. X., NISHIDA, M., SHIMANOUCHI, T. & KUBOI, R. 2008. Liposome membrane itself can affect gene expression in the Escherichia coli cell-free translation system. *Langmuir,* 24**,** 10537-10542.

BULLEID, N. J., BASSEL-DUBY, R. S., FREEDMAN, R. B., SAMBROOK, J. F. & GETHING, M. J. 1992. Cell-free synthesis of enzymically active tissue-type plasminogen activator. Protein folding determines the extent of N-linked glycosylation. *Biochem J,* 286 ( Pt 1)**,** 275-80.

BUNDY, B. C., FRANCISZKOWICZ, M. J. & SWARTZ, J. R. 2008. Escherichia coli-based cell-free synthesis of virus-like particles. *Biotechnol Bioeng,* 100**,** 28-37.

BUNTRU, M., VOGEL, S., SPIEGEL, H. & SCHILLBERG, S. 2014. Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based in vitro translation system. *BMC Biotechnol,* 14**,** 37.

BUNTRU, M., VOGEL, S., STOFF, K., SPIEGEL, H. & SCHILLBERG, S. 2015. A versatile coupled cell-free transcription-translation system based on tobacco BY-2 cell lysates. *Biotechnol Bioeng,* 112**,** 867-78.

CALHOUN, K. A. & SWARTZ, J. R. 2005. Energizing cell-free protein synthesis with glucose metabolism. *Biotechnol Bioeng,* 90**,** 606-13.

CAPPUCCIO, J. A., BLANCHETTE, C. D., SULCHEK, T. A., ARROYO, E. S., KRALJ, J. M., HINZ, A. K., KUHN, E. A., CHROMY, B. A., SEGELKE, B. W., ROTHSCHILD, K. J., FLETCHER, J. E., KATZEN, F., PETERSON, T. C., KUDLICKI, W. A., BENCH, G., HOEPRICH, P. D. & COLEMAN, M. A. 2008. Cell-free co-expression of functional membrane proteins and apolipoprotein, forming soluble nanolipoprotein particles. *Mol Cell Proteomics,* 7**,** 2246-53.

CARLSON, E. D., GAN, R., HODGMAN, C. E. & JEWETT, M. C. 2012. Cell-free protein synthesis: applications come of age. *Biotechnol Adv,* 30**,** 1185-94.

CASTELEIJN, M. G., URTTI, A. & SARKHEL, S. 2013. Expression without boundaries: cell-free protein synthesis in pharmaceutical research. *Int J Pharm,* 440**,** 39-47.

CATHARINE ROSS, A. & ZOLFAGHARI, R. 2011. Cytochrome P450s in the regulation of cellular retinoic acid metabolism. *Annual review of nutrition,* 31**,** 65-87.

CEDERBAUM, A. I. 2015. Molecular mechanisms of the microsomal mixed function oxidases and biological and pathological implications. *Redox Biol,* 4**,** 60-73.

CHEFSON, A. & AUCLAIR, K. 2006. Progress towards the easier use of P450 enzymes. *Molecular BioSystems,* 2**,** 462-469.

CHEN, H. Q., FAN, L. M., XU, Z. N., YIN, X. F. & CEN, P. L. 2007. Efficient production of soluble human beta-defensin-3-4 fusion proteins in Escherichia coli cell-free system. *Process Biochem,* 42**,** 423-428.

CHOI, H.-J., GERMAIN, J. & MONTEMAGNO, C. D. 2006. Effects of different reconstitution procedures on membrane protein activities in proteopolymersomes. *Nanotechnology,* 17**,** 1825.

CHOI, H. J. & MONTEMAGNO, C. D. 2005. Artificial organelle: ATP synthesis from cellular mimetic polymersomes. *Nano Lett,* 5**,** 2538-42.

CHOUDHURY, A., HODGMAN, C. E., ANDERSON, M. J. & JEWETT, M. C. 2014. Evaluating fermentation effects on cell growth and crude extract metabolic activity for improved yeast cell-free protein synthesis. *Biochemical Engineering Journal,* 91**,** 140-148.

CHRISTENSEN, D., KIRBY, D., FOGED, C., AGGER, E. M., ANDERSEN, P., PERRIE, Y. & NIELSEN, H. M. 2008. alpha,alpha'-trehalose 6,6'-dibehenate in non-phospholipid-based liposomes enables direct interaction with trehalose, offering stability during freeze-drying. *Biochim Biophys Acta,* 1778**,** 1365-73.

COLEMAN, W. H. & ROBERTS, W. K. 1981. Factor requirements for the tritin inactivation of animal cell ribosomes. *Biochim Biophys Acta,* 654**,** 57-66.

CORIN, K., BAASKE, P., RAVEL, D. B., SONG, J., BROWN, E., WANG, X., GEISSLER, S., WIENKEN, C. J., JERABEK-WILLEMSEN, M., DUHR, S., BRAUN, D. & ZHANG, S. 2011. A robust and rapid method of producing soluble, stable, and functional G-protein coupled receptors. *PLoS One,* 6**,** e23036.

COSME, J. & JOHNSON, E. F. 2002. Analyzing binding of N-terminal truncated, microsomal cytochrome P450s to membranes. *Methods Enzymol,* 357**,** 116-20.

CRAIG, D., HOWELL, M. T., GIBBS, C. L., HUNT, T. & JACKSON, R. J. 1992. Plasmid cDNA-directed protein synthesis in a coupled eukaryotic in vitro transcription-translation system. *Nucleic Acids Res,* 20**,** 4987-95.

DAVYDOV, D. R. 2011. Microsomal monooxygenase as a multienzyme system: the role of P450-P450 interactions. *Expert opinion on drug metabolism & toxicology,* 7**,** 543-558.

DE HOOG, H. P., LIN JIERONG, E. M., BANERJEE, S., DECAILLOT, F. M. & NALLANI, M. 2014. Conformational antibody binding to a native, cell-free expressed GPCR in block copolymer membranes. *PLoS One,* 9**,** e110847.

DEENI, Y. Y., PAINE, M. J., AYRTON, A. D., CLARKE, S. E., CHENERY, R. & WOLF, C. R. 2001. Expression, purification, and biochemical characterization of a human cytochrome P450 CYP2D6-NADPH cytochrome P450 reductase fusion protein. *Archives of biochemistry and biophysics,* 396**,** 16-24.

DISCHER, B. M., BERMUDEZ, H., HAMMER, D. A., DISCHER, D. E., WON, Y.-Y. & BATES, F. S. 2002. Cross-linked polymersome membranes: vesicles with broadly adjustable properties. *The Journal of Physical Chemistry B,* 106**,** 2848-2854.

DISCHER, B. M., WON, Y. Y., EGE, D. S., LEE, J. C., BATES, F. S., DISCHER, D. E. & HAMMER, D. A. 1999. Polymersomes: tough vesicles made from diblock copolymers. *Science,* 284**,** 1143-6.

DISCHER, D. E. & AHMED, F. 2006. Polymersomes. *Annu Rev Biomed Eng,* 8**,** 323-41.

DISCHER, D. E., ORTIZ, V., SRINIVAS, G., KLEIN, M. L., KIM, Y., CHRISTIAN, D., CAI, S., PHOTOS, P. & AHMED, F. 2007. Emerging Applications of Polymersomes in Delivery: from Molecular Dynamics to Shrinkage of Tumors. *Prog Polym Sci,* 32**,** 838-857.

DOI, N., KUMADAKI, S., OISHI, Y., MATSUMURA, N. & YANAGAWA, H. 2004. In vitro selection of restriction endonucleases by in vitro compartmentalization. *Nucleic Acids Res,* 32**,** e95.

DOMES, S., FILIZ, V., NITSCHE, J., FROMSDORF, A. & FORSTER, S. 2010. Covalent attachment of polymersomes to surfaces. *Langmuir,* 26**,** 6927-31.

DUTHEIL, F., BEAUNE, P. & LORIOT, M. A. 2008. Xenobiotic metabolizing enzymes in the central nervous system: Contribution of cytochrome P450 enzymes in normal and pathological human brain. *Biochimie,* 90**,** 426-36.

EGLI, S., NUSSBAUMER, M. G., BALASUBRAMANIAN, V., CHAMI, M., BRUNS, N., PALIVAN, C. & MEIER, W. 2011a. Biocompatible functionalization of polymersome surfaces: a new approach to surface immobilization and cell targeting using polymersomes. *J Am Chem Soc,* 133**,** 4476-83.

EGLI, S., SCHLAAD, H., BRUNS, N. & MEIER, W. 2011b. Functionalization of block copolymer vesicle surfaces. *Polymers,* 3**,** 252-280.

EKROOS, M. & SJOGREN, T. 2006. Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc Natl Acad Sci U S A,* 103**,** 13682-7.

ENDO, Y., OTSUZUKI, S., ITO, K. & MIURA, K. 1992. Production of an enzymatic active protein using a continuous flow cell-free translation system. *J Biotechnol,* 25**,** 221-30.

ENDO, Y. & SAWASAKI, T. 2006. Cell-free expression systems for eukaryotic protein production. *Curr Opin Biotechnol,* 17**,** 373-80.

EZURE, T., SUZUKI, T. & ANDO, E. 2014. A cell-free protein synthesis system from insect cells. *Methods Mol Biol,* 1118**,** 285-96.

EZURE, T., SUZUKI, T., HIGASHIDE, S., SHINTANI, E., ENDO, K., KOBAYASHI, S., SHIKATA, M., ITO, M., TANIMIZU, K. & NISHIMURA, O. 2006. Cell-free protein synthesis system prepared from insect cells by freeze-thawing. *Biotechnol Prog,* 22**,** 1570-7.

EZURE, T., SUZUKI, T., SHIKATA, M., ITO, M., ANDO, E., NISHIMURA, O. & TSUNASAWA, S. 2007. Expression of proteins containing disulfide bonds in an insect cell-free system and confirmation of their arrangements by MALDI-TOF MS. *Proteomics,* 7**,** 4424-34.

FAROOQ, Y. & ROBERTS, G. C. 2010. Kinetics of electron transfer between NADPH-cytochrome P450 reductase and cytochrome P450 3A4. *Biochem J,* 432**,** 485-93.

FISHER, C. W., CAUDLE, D. L., MARTIN-WIXTROM, C., QUATTROCHI, L. C., TUKEY, R. H., WATERMAN, M. R. & ESTABROOK, R. W. 1992. High-level expression of functional human cytochrome P450 1A2 in Escherichia coli. *FASEB J,* 6**,** 759-64.

FITZGERALD, K. D. & SEMLER, B. L. 2009. Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. *Biochim Biophys Acta,* 1789**,** 518-28.

FRANCO, D., PATHAK, H. B., CAMERON, C. E., ROMBAUT, B., WIMMER, E. & PAUL, A. V. 2005. Stimulation of poliovirus synthesis in a HeLa cell-free in vitro translation-RNA replication system by viral protein 3CDpro. *J Virol,* 79**,** 6358-67.

FRITZ, B. R. & JEWETT, M. C. 2014. The impact of transcriptional tuning on in vitro integrated rRNA transcription and ribosome construction. *Nucleic Acids Res,* 42**,** 6774-85.

FU, Z., OCHSNER, M. A., DE HOOG, H. P., TOMCZAK, N. & NALLANI, M. 2011. Multicompartmentalized polymersomes for selective encapsulation of biomacromolecules. *Chem Commun (Camb),* 47**,** 2862-4.

GAGOSKI, D., MUREEV, S., GILES, N., JOHNSTON, W., DAHMER-HEATH, M., SKALAMERA, D., GONDA, T. J. & ALEXANDROV, K. 2015. Gateway-compatible vectors for high-throughput protein expression in pro- and eukaryotic cell-free systems. *J Biotechnol,* 195**,** 1-7.

GAGOSKI, D., POLINKOVSKY, M. E., MUREEV, S., KUNERT, A., JOHNSTON, W., GAMBIN, Y. & ALEXANDROV, K. 2016. Performance benchmarking of four cell‐free protein expression systems. *Biotechnology and Bioengineering,* 113**,** 292-300.

GALLIE, D. R. 2002. The 5 '-leader of tobacco mosaic virus promotes translation through enhanced recruitment of eIF4F. *Nucleic Acids Research,* 30**,** 3401-3411.

GALLIE, D. R., SLEAT, D. E., WATTS, J. W., TURNER, P. C. & WILSON, T. M. 1987. The 5'-leader sequence of tobacco mosaic virus RNA enhances the expression of foreign gene transcripts in vitro and in vivo. *Nucleic Acids Res,* 15**,** 3257-73.

GAN, R. & JEWETT, M. C. 2014. A combined cell-free transcription-translation system from Saccharomyces cerevisiae for rapid and robust protein synthe. *Biotechnol J,* 9**,** 641-51.

GASPARD, J., CASEY, L. M., ROZIN, M., MUNOZ-PINTO, D. J., SILAS, J. A. & HAHN, M. S. 2016. Mechanical Characterization of Hybrid Vesicles Based on Linear Poly (Dimethylsiloxane-b-Ethylene Oxide) and Poly (Butadiene-b-Ethylene Oxide) Block Copolymers. *Sensors,* 16**,** 390.

GAY, S. C., ROBERTS, A. G. & HALPERT, J. R. 2010. Structural features of cytochromes P450 and ligands that affect drug metabolism as revealed by X-ray crystallography and NMR. *Future Med Chem,* 2**,** 1451-68.

GEORGESCU, J., MUNHOZ, V. H. & BECHINGER, B. 2010. NMR structures of the histidine-rich peptide LAH4 in micellar environments: membrane insertion, pH-dependent mode of antimicrobial action, and DNA transfection. *Biophys J,* 99**,** 2507-15.

GHADESSY, F. J. & HOLLIGER, P. 2004. A novel emulsion mixture for in vitro compartmentalization of transcription and translation in the rabbit reticulocyte system. *Protein Eng Des Sel,* 17**,** 201-4.

GIBBS, P. E., ZOUZIAS, D. C. & FREEDBERG, I. M. 1985. Differential post-translational modification of human type I keratins synthesized in a rabbit reticulocyte cell-free system. *Biochim Biophys Acta,* 824**,** 247-55.

GILEP, A. A., GURYEV, O. L., USANOV, S. A. & ESTABROOK, R. W. 2001. Reconstitution of the Enzymatic Activities of Cytochrome P450s Using Recombinant Flavocytochromes Containing Rat Cytochrome b 5 Fused to NADPH–Cytochrome P450 Reductase with Various Membrane-Binding Segments. *Archives of biochemistry and biophysics,* 390**,** 215-221.

GILLAM, E. M. 2007. Engineering cytochrome P450 enzymes. *Chemical research in toxicology,* 21**,** 220-231.

GILLAM, E. M., BABA, T., KIM, B. R., OHMORI, S. & GUENGERICH, F. P. 1993. Expression of modified human cytochrome P450 3A4 in Escherichia coli and purification and reconstitution of the enzyme. *Arch Biochem Biophys,* 305**,** 123-31.

GILLAM, E. M., GUO, Z., UENG, Y. F., YAMAZAKI, H., COCK, I., REILLY, P. E., HOOPER, W. D. & GUENGERICH, F. P. 1995. Expression of cytochrome P450 3A5 in Escherichia coli: effects of 5' modification, purification, spectral characterization, reconstitution conditions, and catalytic activities. *Arch Biochem Biophys,* 317**,** 374-84.

GIRVAN, H., WALTHAM, T., NEELI, R., COLLINS, H., MCLEAN, K., SCRUTTON, N., LEYS, D. & MUNRO, A. 2006. Flavocytochrome P450 BM3 and the origin of CYP102 fusion species. *Biochemical Society Transactions,* 34**,** 1173-1177.

GLOVER, K. J., WEERAPANA, E. & IMPERIALI, B. 2005a. In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc Natl Acad Sci U S A,* 102**,** 14255-9.

GLOVER, K. J., WEERAPANA, E., NUMAO, S. & IMPERIALI, B. 2005b. Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from Campylobacter jejuni. *Chem Biol,* 12**,** 1311-5.

GOREN, M. A. & FOX, B. G. 2008. Wheat germ cell-free translation, purification, and assembly of a functional human stearoyl-CoA desaturase complex. *Protein Expr Purif,* 62**,** 171-8.

GOSHIMA, N., KAWAMURA, Y., FUKUMOTO, A., MIURA, A., HONMA, R., SATOH, R., WAKAMATSU, A., YAMAMOTO, J., KIMURA, K., NISHIKAWA, T., ANDOH, T., IIDA, Y., ISHIKAWA, K., ITO, E., KAGAWA, N., KAMINAGA, C., KANEHORI, K., KAWAKAMI, B., KENMOCHI, K., KIMURA, R., KOBAYASHI, M., KUROITA, T., KUWAYAMA, H., MARUYAMA, Y., MATSUO, K., MINAMI, K., MITSUBORI, M., MORI, M., MORISHITA, R., MURASE, A., NISHIKAWA, A., NISHIKAWA, S., OKAMOTO, T., SAKAGAMI, N., SAKAMOTO, Y., SASAKI, Y., SEKI, T., SONO, S., SUGIYAMA, A., SUMIYA, T., TAKAYAMA, T., TAKAYAMA, Y., TAKEDA, H., TOGASHI, T., YAHATA, K., YAMADA, H., YANAGISAWA, Y., ENDO, Y., IMAMOTO, F., KISU, Y., TANAKA, S., ISOGAI, T., IMAI, J., WATANABE, S. & NOMURA, N. 2008. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat Methods,* 5**,** 1011-7.

GRIFFITHS, A. D. & TAWFIK, D. S. 2003. Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization. *EMBO J,* 22**,** 24-35.

GRINKOVA, Y. V., DENISOV, I. G. & SLIGAR, S. G. 2010. Functional reconstitution of monomeric CYP3A4 with multiple cytochrome P450 reductase molecules in Nanodiscs. *Biochem Biophys Res Commun,* 398**,** 194-8.

GUARINO, C. & DELISA, M. P. 2012. A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins. *Glycobiology,* 22**,** 596-601.

GUDKOV, A. T., OZEROVA, M. V., SHIRYAEV, V. M. & SPIRIN, A. S. 2005. 5 '-poly(A) sequence as an effective leader for translation in eukaryotic cell-free systems. *Biotechnology and Bioengineering,* 91**,** 468-473.

GUENGERICH, F. P. 1999. Cytochrome P-450 3A4: regulation and role in drug metabolism. *Annu Rev Pharmacol Toxicol,* 39**,** 1-17.

GURYEV, O. L., GILEP, A. A., USANOV, S. A. & ESTABROOK, R. W. 2001. Interaction of apo-cytochrome b5 with cytochromes P4503A4 and P45017A: relevance of heme transfer reactions. *Biochemistry,* 40**,** 5018-31.

HANES, J. & PLUCKTHUN, A. 1997. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A,* 94**,** 4937-42.

HE, M. & TAUSSIG, M. J. 2002. Ribosome display: cell-free protein display technology. *Brief Funct Genomic Proteomic,* 1**,** 204-12.

HEIN, C., HENRICH, E., ORBAN, E., DOTSCH, V. & BERNHARD, F. 2014. Hydrophobic supplements in cell-free systems: Designing artificial environments for membrane proteins. *Engineering in Life Sciences,* 14**,** 365-379.

HELENIUS, A. & AEBI, M. 2004. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem,* 73**,** 1019-49.

HOCKENBERRY, A. J. & JEWETT, M. C. 2012. Synthetic in vitro circuits. *Curr Opin Chem Biol,* 16**,** 253-9.

HODGMAN, C. E. & JEWETT, M. C. 2012. Cell-free synthetic biology: thinking outside the cell. *Metab Eng,* 14**,** 261-9.

HODGMAN, C. E. & JEWETT, M. C. 2013. Optimized extract preparation methods and reaction conditions for improved yeast cell-free protein synthesis. *Biotechnol Bioeng,* 110**,** 2643-54.

HODGMAN, C. E. & JEWETT, M. C. 2014. Characterizing IGR IRES-mediated translation initiation for use in yeast cell-free protein synthesis. *New Biotechnology,* 31**,** 499-505.

HOHSAKA, T., ASHIZUKA, Y., MURAKAMI, H. & SISIDO, M. 2001a. Five-base codons for incorporation of nonnatural amino acids into proteins. *Nucleic Acids Res,* 29**,** 3646-51.

HOHSAKA, T., ASHIZUKA, Y., TAIRA, H., MURAKAMI, H. & SISIDO, M. 2001b. Incorporation of nonnatural amino acids into proteins by using various four-base codons in an Escherichia coli in vitro translation system. *Biochemistry,* 40**,** 11060-4.

HOLMANS, P. L., SHET, M. S., MARTIN-WIXTROM, C. A., FISHER, C. W. & ESTABROOK, R. W. 1994. The high-level expression in Escherichia coli of the membrane-bound form of human and rat cytochrome b5 and studies on their mechanism of function. *Arch Biochem Biophys,* 312**,** 554-65.

HOVIJITRA, N. T., WUU, J. J., PEAKER, B. & SWARTZ, J. R. 2009. Cell-free synthesis of functional aquaporin Z in synthetic liposomes. *Biotechnol Bioeng,* 104**,** 40-9.

IGARASHI, K., KOJIMA, M., WATANABE, Y., MAEDA, K. & HIROSE, S. 1980. Stimulation of Polypeptide-Synthesis by Spermidine at the Level of Initiation in Rabbit Reticulocyte and Wheat-Germ Cell-Free Systems. *Biochemical and Biophysical Research Communications,* 97**,** 480-486.

ITEL, F., CHAMI, M., NAJER, A., LÖRCHER, S., WU, D., DINU, I. A. & MEIER, W. 2014. Molecular organization and dynamics in polymersome membranes: A lateral diffusion study. *Macromolecules,* 47**,** 7588-7596.

JEWETT, M. C. & FORSTER, A. C. 2010. Update on designing and building minimal cells. *Curr Opin Biotechnol,* 21**,** 697-703.

JEWETT, M. C., FRITZ, B. R., TIMMERMAN, L. E. & CHURCH, G. M. 2013. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol Syst Biol,* 9**,** 678.

JOSHI, B., CAI, A. L., KEIPER, B. D., MINICH, W. B., MENDEZ, R., BEACH, C. M., STEPINSKI, J., STOLARSKI, R., DARZYNKIEWICZ, E. & RHOADS, R. E. 1995. Phosphorylation of eukaryotic protein synthesis initiation factor 4E at Ser-209. *J Biol Chem,* 270**,** 14597-603.

KAHN, J. S., RUIZ, R. C., SUREKA, S., PENG, S., DERRIEN, T. L., AN, D. & LUO, D. 2016. DNA Microgels as a Platform for Cell-free Protein Expression and Display. *Biomacromolecules*.

KALMBACH, R., CHIZHOV, I., SCHUMACHER, M. C., FRIEDRICH, T., BAMBERG, E. & ENGELHARD, M. 2007. Functional cell-free synthesis of a seven helix membrane protein: in situ insertion of bacteriorhodopsin into liposomes. *J Mol Biol,* 371**,** 639-48.

KAMURA, N., SAWASAKI, T., KASAHARA, Y., TAKAI, K. & ENDO, Y. 2005. Selection of 5'-untranslated sequences that enhance initiation of translation in a cell-free protein synthesis system from wheat embryos. *Bioorg Med Chem Lett,* 15**,** 5402-6.

KANDA, T., TAKAI, K., HOHSAKA, T., SISIDO, M. & TAKAKU, H. 2000. Sense codon-dependent introduction of unnatural amino acids into multiple sites of a protein. *Biochem Biophys Res Commun,* 270**,** 1136-9.

KANEDA, M., NOMURA, S. M., ICHINOSE, S., KONDO, S., NAKAHAMA, K., AKIYOSHI, K. & MORITA, I. 2009. Direct formation of proteo-liposomes by in vitro synthesis and cellular cytosolic delivery with connexin-expressing liposomes. *Biomaterials,* 30**,** 3971-7.

KANNO, T. & TOZAWA, Y. 2010. Protein engineering accelerated by cell-free technology. *Methods Mol Biol,* 607**,** 85-99.

KANTER, G., YANG, J., VOLOSHIN, A., LEVY, S., SWARTZ, J. R. & LEVY, R. 2007. Cell-free production of scFv fusion proteins: an efficient approach for personalized lymphoma vaccines. *Blood,* 109**,** 3393-9.

KATZ, J. S., ZHONG, S., RICART, B. G., POCHAN, D. J., HAMMER, D. A. & BURDICK, J. A. 2010. Modular synthesis of biodegradable diblock copolymers for designing functional polymersomes. *J Am Chem Soc,* 132**,** 3654-5.

KATZEN, F., CHANG, G. & KUDLICKI, W. 2005. The past, present and future of cell-free protein synthesis. *Trends Biotechnol,* 23**,** 150-6.

KATZEN, F. & KUDLICKI, W. 2006. Efficient generation of insect-based cell-free translation extracts active in glycosylation and signal sequence processing. *J Biotechnol,* 125**,** 194-7.

KAWASAKI, T., GOUDA, M. D., SAWASAKI, T., TAKAI, K. & ENDO, Y. 2003. Efficient synthesis of a disulfide-containing protein through a batch cell-free system from wheat germ. *Eur J Biochem,* 270**,** 4780-6.

KIGAWA, T., YABUKI, T., MATSUDA, N., MATSUDA, T., NAKAJIMA, R., TANAKA, A. & YOKOYAMA, S. 2004. Preparation of Escherichia coli cell extract for highly productive cell-free protein expression. *Journal of structural and functional genomics,* 5**,** 63-68.

KIGAWA, T. & YOKOYAMA, S. 1991. A continuous cell-free protein synthesis system for coupled transcription-translation. *J Biochem,* 110**,** 166-8.

KIM, D. & GUENGERICH, F. P. 2004. Enhancement of 7-methoxyresorufin O-demethylation activity of human cytochrome P450 1A2 by molecular breeding. *Archives of biochemistry and biophysics,* 432**,** 102-108.

KIM, D. M. & SWARTZ, J. R. 1999. Prolonging cell-free protein synthesis with a novel ATP regeneration system. *Biotechnol Bioeng,* 66**,** 180-8.

KIM, D. M. & SWARTZ, J. R. 2000. Prolonging cell-free protein synthesis by selective reagent additions. *Biotechnol Prog,* 16**,** 385-90.

KIM, D. M. & SWARTZ, J. R. 2004. Efficient production of a bioactive, multiple disulfide-bonded protein using modified extracts of Escherichia coli. *Biotechnol Bioeng,* 85**,** 122-9.

KIM, H. C., KIM, T. W. & KIM, D. M. 2011. Prolonged production of proteins in a cell-free protein synthesis system using polymeric carbohydrates as an energy source. *Process Biochemistry,* 46**,** 1366-1369.

KIM, J. & WINFREE, E. 2011. Synthetic in vitro transcriptional oscillators. *Mol Syst Biol,* 7**,** 465.

KIM, T. W., KEUM, J. W., OH, I. S., CHOI, C. Y., KIM, H. C. & KIM, D. M. 2007. An economical and highly productive cell-free protein synthesis system utilizing fructose-1,6-bisphosphate as an energy source. *J Biotechnol,* 130**,** 389-93.

KLEIN, K. & ZANGER, U. M. 2013. Pharmacogenomics of Cytochrome P450 3A4: Recent Progress Toward the "Missing Heritability" Problem. *Front Genet,* 4**,** 12.

KLERMUND, L., POSCHENRIEDER, S. T. & CASTIGLIONE, K. 2016. Simple surface functionalization of polymersomes using non-antibacterial peptide anchors. *J Nanobiotechnology,* 14**,** 48.

KNAPP, K. G., GOERKE, A. R. & SWARTZ, J. R. 2007. Cell-free synthesis of proteins that require disulfide bonds using glucose as an energy source. *Biotechnol Bioeng,* 97**,** 901-8.

KOBAYASHI, T., MACHIDA, K., MIKAMI, S., MASUTANI, M. & IMATAKA, H. 2011. Cell-free RNA replication systems based on a human cell extracts-derived in vitro translation system with the encephalomyocarditisvirus RNA. *J Biochem,* 150**,** 423-30.

KOBAYASHI, T., NAKAMURA, Y., MIKAMI, S., MASUTANI, M., MACHIDA, K. & IMATAKA, H. 2012. Synthesis of encephalomyocarditis virus in a cell-free system: from DNA to RNA virus in one tube. *Biotechnol Lett,* 34**,** 67-73.

KOWAL, J. L., KOWAL, J. K., WU, D., STAHLBERG, H., PALIVAN, C. G. & MEIER, W. P. 2014. Functional surface engineering by nucleotide-modulated potassium channel insertion into polymer membranes attached to solid supports. *Biomaterials,* 35**,** 7286-94.

KRINSKY, N., KADURI, M., SHAINSKY-ROITMAN, J., GOLDFEDER, M., IVANIR, E., BENHAR, I., SHOHAM, Y. & SCHROEDER, A. 2016. A Simple and Rapid Method for Preparing a Cell-Free Bacterial Lysate for Protein Synthesis. *PloS one,* 11**,** e0165137.

KUMAR, M., GRZELAKOWSKI, M., ZILLES, J., CLARK, M. & MEIER, W. 2007. Highly permeable polymeric membranes based on the incorporation of the functional water channel protein Aquaporin Z. *Proc Natl Acad Sci U S A,* 104**,** 20719-24.

KUMAR, S., CHEN, C. S., WAXMAN, D. J. & HALPERT, J. R. 2005. Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. *J Biol Chem,* 280**,** 19569-75.

KUMAR, S., LIU, H. & HALPERT, J. R. 2006. Engineering of cytochrome P450 3A4 for enhanced peroxide-mediated substrate oxidation using directed evolution and site-directed mutagenesis. *Drug metabolism and disposition,* 34**,** 1958-1965.

LACROIX, D., SONNIER, M., MONCION, A., CHERON, G. & CRESTEIL, T. 1997. Expression of CYP3A in the human liver--evidence that the shift between CYP3A7 and CYP3A4 occurs immediately after birth. *Eur J Biochem,* 247**,** 625-34.

LAMB, D. C., KELLY, D. E., BALDWIN, B. C. & KELLY, S. L. 2000. Differential inhibition of human CYP3A4 and Candida albicans CYP51 with azole antifungal agents. *Chem Biol Interact,* 125**,** 165-75.

LARSON, J. R., COON, M. J. & PORTER, T. D. 1991. Purification and properties of a shortened form of cytochrome P-450 2E1: deletion of the NH2-terminal membrane-insertion signal peptide does not alter the catalytic activities. *Proc Natl Acad Sci U S A,* 88**,** 9141-5.

LECLERCQ, L. & VERT, M. 2017. Comparison between protein repulsions by diblock PLA-PEO and albumin nanocoatings using OWLS. *J Biomater Sci Polym Ed,* 28**,** 177-193.

LEE, C. A., KADWELL, S. H., KOST, T. A. & SERABJIT-SINGH, C. J. 1995. CYP3A4 expressed by insect cells infected with a recombinant baculovirus containing both CYP3A4 and human NADPH-cytochrome P450 reductase is catalytically similar to human liver microsomal CYP3A4. *Arch Biochem Biophys,* 319**,** 157-67.

LEE, H., KIM, J.-H., HAN, S., LIM, Y.-R., PARK, H.-G., CHUN, Y.-J., PARK, S.-W. & KIM, D. 2014. Directed-evolution analysis of human cytochrome P450 2A6 for enhanced enzymatic catalysis. *Journal of Toxicology and Environmental Health, Part A,* 77**,** 1409-1418.

LEE, J. C. M., BERMUDEZ, H., DISCHER, B. M., SHEEHAN, M. A., WON, Y. Y., BATES, F. S. & DISCHER, D. E. 2001. Preparation, stability, and in vitro performance of vesicles made with diblock copolymers. *Biotechnology and bioengineering,* 73**,** 135-145.

LEE, J. S., CHEONG, H. S., KIM, L. H., KIM, J. O., SEO, D. W., KIM, Y. H., CHUNG, M. W., HAN, S. Y. & SHIN, H. D. 2013a. Screening of Genetic Polymorphisms of CYP3A4 and CYP3A5 Genes. *Korean J Physiol Pharmacol,* 17**,** 479-84.

LEE, M. T., SUN, T. L., HUNG, W. C. & HUANG, H. W. 2013b. Process of inducing pores in membranes by melittin. *Proc Natl Acad Sci U S A,* 110**,** 14243-8.

LEE, S. J. & GOLDSTEIN, J. A. 2012. Comparison of CYP3A4 and CYP3A5: the effects of cytochrome b5 and NADPH-cytochrome P450 reductase on testosterone hydroxylation activities. *Drug Metab Pharmacokinet,* 27**,** 663-7.

LEVERITT, J. M., 3RD, PINO-ANGELES, A. & LAZARIDIS, T. 2015. The structure of a melittin-stabilized pore. *Biophys J,* 108**,** 2424-6.

LI, A. P., KAMINSKI, D. L. & RASMUSSEN, A. 1995. Substrates of human hepatic cytochrome P450 3A4. *Toxicology,* 104**,** 1-8.

LI, S., BYRNE, B., WELSH, J. & PALMER, A. F. 2007. Self-assembled poly(butadiene)-b-poly(ethylene oxide) polymersomes as paclitaxel carriers. *Biotechnol Prog,* 23**,** 278-85.

LIAN, J., MA, Y., CAI, J., WU, M., WANG, J., WANG, X. & XU, Z. 2009. High-level expression of soluble subunit b of F1F0 ATP synthase in Escherichia coli cell-free system. *Appl Microbiol Biotechnol,* 85**,** 303-11.

LIGUORI, L., BLESNEAC, I., MADERN, D., VIVAUDOU, M. & LENORMAND, J. L. 2010. Single-step production of functional OEP24 proteoliposomes. *Protein Expr Purif,* 69**,** 106-11.

LIGUORI, L. & LENORMAND, J. L. 2009. Production of recombinant proteoliposomes for therapeutic uses. *Methods Enzymol,* 465**,** 209-23.

LIM, S. K., DE HOOG, H.-P., PARIKH, A. N., NALLANI, M. & LIEDBERG, B. 2013. Hybrid, nanoscale phospholipid/block copolymer vesicles. *Polymers,* 5**,** 1102-1114.

LIN, J. J., SILAS, J. A., BERMUDEZ, H., MILAM, V. T., BATES, F. S. & HAMMER, D. A. 2004. The effect of polymer chain length and surface density on the adhesiveness of functionalized polymersomes. *Langmuir,* 20**,** 5493-500.

LINGAPPA, V. R., LINGAPPA, J. R., PRASAD, R., EBNER, K. E. & BLOBEL, G. 1978. Coupled cell-free synthesis, segregation, and core glycosylation of a secretory protein. *Proc Natl Acad Sci U S A,* 75**,** 2338-42.

LIU, D. V., ZAWADA, J. F. & SWARTZ, J. R. 2005. Streamlining Escherichia coli S30 extract preparation for economical cell-free protein synthesis. *Biotechnol Prog,* 21**,** 460-5.

LOCUSON, C. W., WIENKERS, L. C., JONES, J. P. & TRACY, T. S. 2007. CYP2C9 protein interactions with cytochrome b(5): effects on the coupling of catalysis. *Drug Metab Dispos,* 35**,** 1174-81.

LYUKMANOVA, E. N., SHENKAREV, Z. O., KHABIBULLINA, N. F., KOPEINA, G. S., SHULEPKO, M. A., PARAMONOV, A. S., MINEEV, K. S., TIKHONOV, R. V., SHINGAROVA, L. N., PETROVSKAYA, L. E., DOLGIKH, D. A., ARSENIEV, A. S. & KIRPICHNIKOV, M. P. 2012. Lipid-protein nanodiscs for cell-free production of integral membrane proteins in a soluble and folded state: comparison with detergent micelles, bicelles and liposomes. *Biochim Biophys Acta,* 1818**,** 349-58.

MADIN, K., SAWASAKI, T., OGASAWARA, T. & ENDO, Y. 2000. A highly efficient and robust cell-free protein synthesis system prepared from wheat embryos: plants apparently contain a suicide system directed at ribosomes. *Proc Natl Acad Sci U S A,* 97**,** 559-64.

MARCUS, A. & FEELEY, J. 1964. Activation of Protein Synthesis in the Imbibition Phase of Seed Germination. *Proc Natl Acad Sci U S A,* 51**,** 1075-9.

MARTINO, C., KIM, S. H., HORSFALL, L., ABBASPOURRAD, A., ROSSER, S. J., COOPER, J. & WEITZ, D. A. 2012. Protein expression, aggregation, and triggered release from polymersomes as artificial cell-like structures. *Angew Chem Int Ed Engl,* 51**,** 6416-20.

MASUDA, T., GOTO, F., YOSHIHARA, T., EZURE, T., SUZUKI, T., KOBAYASHI, S., SHIKATA, M. & UTSUMI, S. 2007. Construction of homo- and heteropolymers of plant ferritin subunits using an in vitro protein expression system. *Protein Expr Purif,* 56**,** 237-46.

MAY, S., ANDREASSON-OCHSNER, M., FU, Z., LOW, Y. X., TAN, D., DE HOOG, H. P., RITZ, S., NALLANI, M. & SINNER, E. K. 2013. In vitro expressed GPCR inserted in polymersome membranes for ligand-binding studies. *Angew Chem Int Ed Engl,* 52**,** 749-53.

MEIER, W., NARDIN, C. & WINTERHALTER, M. 2000. Reconstitution of Channel Proteins in (Polymerized) ABA Triblock Copolymer Membranes. *Angew Chem Int Ed Engl,* 39**,** 4599-4602.

MESSAGER, L., GAITZSCH, J., CHIERICO, L. & BATTAGLIA, G. 2014. Novel aspects of encapsulation and delivery using polymersomes. *Curr Opin Pharmacol,* 18**,** 104-11.

MICHEL, E. & WUTHRICH, K. 2012. High-yield Escherichia coli-based cell-free expression of human proteins. *J Biomol NMR,* 53**,** 43-51.

MICHEL, Y. M., PONCET, D., PIRON, M., KEAN, K. M. & BORMAN, A. M. 2000. Cap-Poly(A) synergy in mammalian cell-free extracts. Investigation of the requirements for poly(A)-mediated stimulation of translation initiation. *J Biol Chem,* 275**,** 32268-76.

MIKAMI, S., KOBAYASHI, T., YOKOYAMA, S. & IMATAKA, H. 2006a. A hybridoma-based in vitro translation system that efficiently synthesizes glycoproteins. *J Biotechnol,* 127**,** 65-78.

MIKAMI, S., MASUTANI, M., SONENBERG, N., YOKOYAMA, S. & IMATAKA, H. 2006b. An efficient mammalian cell-free translation system supplemented with translation factors. *Protein Expr Purif,* 46**,** 348-57.

MOLLA, A., PAUL, A. V. & WIMMER, E. 1991. Cell-free, de novo synthesis of poliovirus. *Science,* 254**,** 1647-51.

MORITANI, Y., NOMURA, S. M., MORITA, I. & AKIYOSHI, K. 2010. Direct integration of cell-free-synthesized connexin-43 into liposomes and hemichannel formation. *Febs Journal,* 277**,** 3343-3352.

MOULAY, G., LEBORGNE, C., MASON, A. J., AISENBREY, C., KICHLER, A. & BECHINGER, B. 2017. Histidine-rich designer peptides of the LAH4 family promote cell delivery of a multitude of cargo. *J Pept Sci,* 23**,** 320-328.

MUNRO, A. W., LEYS, D. G., MCLEAN, K. J., MARSHALL, K. R., OST, T. W. B., DAFF, S., MILES, C. S., CHAPMAN, S. K., LYSEK, D. A., MOSER, C. C., PAGE, C. C. & DUTTON, P. L. 2002. P450 BM3: The very model of a modern flavocytochrome. *Trends in Biochemical Sciences,* 27**,** 250-257.

MURANAKA, N. & YOKOBAYASHI, Y. 2010. Posttranscriptional signal integration of engineered riboswitches yields band-pass output. *Angew Chem Int Ed Engl,* 49**,** 4653-5.

MUREEV, S., KOVTUN, O., NGUYEN, U. T. & ALEXANDROV, K. 2009. Species-independent translational leaders facilitate cell-free expression. *Nat Biotechnol,* 27**,** 747-52.

MURRAY, C. J. & BALIGA, R. 2013. Cell-free translation of peptides and proteins: from high throughput screening to clinical production. *Curr Opin Chem Biol,* 17**,** 420-6.

NALLANI, M., ANDREASSON-OCHSNER, M., TAN, C. W., SINNER, E. K., WISANTOSO, Y., GEIFMAN-SHOCHAT, S. & HUNZIKER, W. 2011. Proteopolymersomes: in vitro production of a membrane protein in polymersome membranes. *Biointerphases,* 6**,** 153-7.

NALLANI, M., WOESTENENK, R., DE HOOG, H. P., VAN DONGEN, S. F., BOEZEMAN, J., CORNELISSEN, J. J., NOLTE, R. J. & VAN HEST, J. C. 2009. Sorting catalytically active polymersome nanoreactors by flow cytometry. *Small,* 5**,** 1138-43.

NARDIN, C., HIRT, T., LEUKEL, J. & MEIER, W. 2000. Polymerized ABA triblock copolymer vesicles. *Langmuir,* 16**,** 1035-1041.

NEUMANN, H. 2012. Rewiring translation - Genetic code expansion and its applications. *FEBS Lett,* 586**,** 2057-64.

NOIREAUX, V., BAR-ZIV, R. & LIBCHABER, A. 2003. Principles of cell-free genetic circuit assembly. *Proc Natl Acad Sci U S A,* 100**,** 12672-7.

NOMURA, S. M., KONDOH, S., ASAYAMA, W., ASADA, A., NISHIKAWA, S. & AKIYOSHI, K. 2008. Direct preparation of giant proteo-liposomes by in vitro membrane protein synthesis. *J Biotechnol,* 133**,** 190-5.

NOOR, M., DWORECK, T., SCHENK, A., SHINDE, P., FIORONI, M. & SCHWANEBERG, U. 2012. Polymersome surface decoration by an EGFP fusion protein employing Cecropin A as peptide "anchor". *J Biotechnol,* 157**,** 31-7.

NOREN, C. J., ANTHONY-CAHILL, S. J., GRIFFITH, M. C. & SCHULTZ, P. G. 1989. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science,* 244**,** 182-8.

NOURIAN, Z., ROELOFSEN, W. & DANELON, C. 2012. Triggered gene expression in fed-vesicle microreactors with a multifunctional membrane. *Angew Chem Int Ed Engl,* 51**,** 3114-8.

OGAWA, A., TABUCHI, J. & DOI, Y. 2014. Identification of short untranslated regions that sufficiently enhance translation in high-quality wheat germ extract. *Bioorg Med Chem Lett,* 24**,** 3724-7.

OH, I.-S., KIM, T.-W., AHN, J.-H., KEUM, J.-W., CHOI, C.-Y. & KIM, D.-M. 2007. Use of l-buthionine sulfoximine for the efficient expression of disulfide-containing proteins in cell-free extracts of Escherichia coli. *Biotechnol Bioprocess Eng,* 12**,** 574-8.

OZAWA, K., LOSCHA, K. V., KUPPAN, K. V., LOH, C. T., DIXON, N. E. & OTTING, G. 2012. High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. *Biochem Biophys Res Commun,* 418**,** 652-6.

PALMER, E., LIU, H., KHAN, F., TAUSSIG, M. J. & HE, M. 2006. Enhanced cell-free protein expression by fusion with immunoglobulin Ckappa domain. *Protein Sci,* 15**,** 2842-6.

PARK, N., UM, S. H., FUNABASHI, H., XU, J. & LUO, D. 2009. A cell-free protein-producing gel. *Nat Mater,* 8**,** 432-7.

PATEL, K. G. & SWARTZ, J. R. 2011. Surface functionalization of virus-like particles by direct conjugation using azide-alkyne click chemistry. *Bioconjug Chem,* 22**,** 376-87.

PAULUS, M., HASLBECK, M. & WATZELE, M. 2004. RNA stem-loop enhanced expression of previously non-expressible genes. *Nucleic Acids Res,* 32**,** e78.

PELHAM, H. R. & JACKSON, R. J. 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. *Eur J Biochem,* 67**,** 247-56.

PELLETIER, J. & SONENBERG, N. 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature,* 334**,** 320-5.

PERNECKY, S. J., LARSON, J. R., PHILPOT, R. M. & COON, M. J. 1993. Expression of truncated forms of liver microsomal P450 cytochromes 2B4 and 2E1 in Escherichia coli: influence of NH2-terminal region on localization in cytosol and membranes. *Proc Natl Acad Sci U S A,* 90**,** 2651-5.

PERNECKY, S. J., OLKEN, N. M., BESTERVELT, L. L. & COON, M. J. 1995. Subcellular localization, aggregation state, and catalytic activity of microsomal P450 cytochromes modified in the NH2-terminal region and expressed in Escherichia coli. *Arch Biochem Biophys,* 318**,** 446-56.

PERRET, A. & POMPON, D. 1998. Electron shuttle between membrane-bound cytochrome P450 3A4 and b5 rules uncoupling mechanisms. *Biochemistry,* 37**,** 11412-24.

PERRONE, B., MILES, A. J., SALNIKOV, E. S., WALLACE, B. A. & BECHINGER, B. 2014. Lipid interactions of LAH4, a peptide with antimicrobial and nucleic acid transfection activities. *Eur Biophys J,* 43**,** 499-507.

PINTO, J. T. & COOPER, A. J. 2014. From cholesterogenesis to steroidogenesis: role of riboflavin and flavoenzymes in the biosynthesis of vitamin D. *Advances in Nutrition: An International Review Journal,* 5**,** 144-163.

PORTER, T. D. 2002. The roles of cytochrome b5 in cytochrome P450 reactions. *J Biochem Mol Toxicol,* 16**,** 311-6.

PROSSER, D. E. & JONES, G. 2004. Enzymes involved in the activation and inactivation of vitamin D. *Trends in biochemical sciences,* 29**,** 664-673.

QUAST, R. B., KORTT, O., HENKEL, J., DONDAPATI, S. K., WUSTENHAGEN, D. A., STECH, M. & KUBICK, S. 2015. Automated production of functional membrane proteins using eukaryotic cell-free translation systems. *Journal of Biotechnology,* 203**,** 45-53.

RAJESH, S., KNOWLES, T. & OVERDUIN, M. 2011. Production of membrane proteins without cells or detergents. *New Biotechnology,* 28**,** 250-254.

RAMEEZ, S., ALOSTA, H. & PALMER, A. F. 2008. Biocompatible and biodegradable polymersome encapsulated hemoglobin: a potential oxygen carrier. *Bioconjug Chem,* 19**,** 1025-32.

RENDIC, S. & GUENGERICH, F. P. 2015. Survey of Human Oxidoreductases and Cytochrome P450 Enzymes Involved in the Metabolism of Xenobiotic and Natural Chemicals. *Chem Res Toxicol,* 28**,** 38-42.

RICHTER, C., BICKEL, F., OSBERGHAUS, A. & HUBBUCH, J. 2014. High‐throughput characterization of an insect cell‐free expression. *Engineering in Life Sciences*.

RIGAUD, J. L. & LEVY, D. 2003. Reconstitution of membrane proteins into liposomes. *Methods Enzymol,* 372**,** 65-86.

ROBERTS, G. A., GROGAN, G., GRETER, A., FLITSCH, S. L. & TURNER, N. J. 2002. Identification of a New Class of Cytochrome P450 from a Rhodococcus sp. *Journal of bacteriology,* 184**,** 3898-3908.

RODRÍGUEZ-GARCÍA, R., MELL, M., LÓPEZ-MONTERO, I., NETZEL, J., HELLWEG, T. & MONROY, F. 2011. Polymersomes: smart vesicles of tunable rigidity and permeability. *Soft Matter,* 7**,** 1532-1542.

RODRÍGUEZ-LIMAS, W. A., SEKAR, K. & TYO, K. E. J. 2013. Virus-like particles: the future of microbial factories and cell-free systems as platforms for vaccine development. *Current Opinion in Biotechnology,* 24**,** 1089-1093.

RUBENSTEIN, P., SMITH, P., DEUCHLER, J. & REDMAN, K. 1981. NH2-terminal acetylation of Dictyostelium discoideum actin in a cell-free protein-synthesizing system. *J Biol Chem,* 256**,** 8149-55.

RYABOVA, L. A., DESPLANCQ, D., SPIRIN, A. S. & PLUCKTHUN, A. 1997. Functional antibody production using cell-free translation: effects of protein disulfide isomerase and chaperones. *Nat Biotechnol,* 15**,** 79-84.

RYABOVA, L. A., TORGASHOV, A. F., KURNASOV, O. V., BUBUNENKO, M. G. & SPIRIN, A. S. 1993. The 3'-terminal untranslated region of alfalfa mosaic virus RNA 4 facilitates the RNA entry into translation in a cell-free system. *FEBS Lett,* 326**,** 264-6.

SACHSE, R., DONDAPATI, S. K., FENZ, S. F., SCHMIDT, T. & KUBICK, S. 2014. Membrane protein synthesis in cell-free systems: from bio-mimetic systems to bio-membranes. *FEBS Lett,* 588**,** 2774-81.

SAKAGUCHI, M., MIHARA, K. & SATO, R. 1984. Signal recognition particle is required for co-translational insertion of cytochrome P-450 into microsomal membranes. *Proc Natl Acad Sci U S A,* 81**,** 3361-4.

SAKAGUCHI, M., MIHARA, K. & SATO, R. 1987. A short amino-terminal segment of microsomal cytochrome P-450 functions both as an insertion signal and as a stop-transfer sequence. *EMBO J,* 6**,** 2425-31.

SANDHU, P., GUO, Z., BABA, T., MARTIN, M. V., TUKEY, R. H. & GUENGERICH, F. P. 1994. Expression of modified human cytochrome P450 1A2 in Escherichia coli: stabilization, purification, spectral characterization, and catalytic activities of the enzyme. *Arch Biochem Biophys,* 309**,** 168-77.

SANFORD, J., CODINA, J. & BIRNBAUMER, L. 1991. Gamma-subunits of G proteins, but not their alpha- or beta-subunits, are polyisoprenylated. Studies on post-translational modifications using in vitro translation with rabbit reticulocyte lysates. *J Biol Chem,* 266**,** 9570-9.

SATO, H. & FEIX, J. B. 2006. Peptide-membrane interactions and mechanisms of membrane destruction by amphipathic alpha-helical antimicrobial peptides. *Biochim Biophys Acta,* 1758**,** 1245-56.

SCHENKMAN, J. B. & JANSSON, I. 2003. The many roles of cytochrome b5. *Pharmacol Ther,* 97**,** 139-52.

SCHIFFLER, B. & BERNHARDT, R. 2003. Bacterial (CYP101) and mitochondrial P450 systems—how comparable are they? *Biochemical and biophysical research communications,* 312**,** 223-228.

SCHWARZ, D., DOTSCH, V. & BERNHARD, F. 2008. Production of membrane proteins using cell-free expression systems. *Proteomics,* 8**,** 3933-46.

SCOTT, E. E., HE, Y. A., WESTER, M. R., WHITE, M. A., CHIN, C. C., HALPERT, J. R., JOHNSON, E. F. & STOUT, C. D. 2003. An open conformation of mammalian cytochrome P450 2B4 at 1.6-A resolution. *Proc Natl Acad Sci U S A,* 100**,** 13196-201.

SEDDON, A. M., CURNOW, P. & BOOTH, P. J. 2004. Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta,* 1666**,** 105-17.

SEVRIOUKOVA, I. F. & POULOS, T. L. 2017. Structural basis for regiospecific midazolam oxidation by human cytochrome P450 3A4. *Proc Natl Acad Sci U S A,* 114**,** 486-491.

SHIMIZU, Y., INOUE, A., TOMARI, Y., SUZUKI, T., YOKOGAWA, T., NISHIKAWA, K. & UEDA, T. 2001. Cell-free translation reconstituted with purified components. *Nat Biotechnol,* 19**,** 751-5.

SHIN, J. & NOIREAUX, V. 2012. An E. coli cell-free expression toolbox: application to synthetic gene circuits and artificial cells. *ACS Synth Biol,* 1**,** 29-41.

SHRESTHA, P., HOLLAND, T. M. & BUNDY, B. C. 2012. Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques,* 53**,** 163-74.

SMITH, G. E., SUMMERS, M. D. & FRASER, M. J. 1983. Production of human beta interferon in insect cells infected with a baculovirus expression vector. *Mol Cell Biol,* 3**,** 2156-65.

SMITH, M. T., VARNER, C. T., BUSH, D. B. & BUNDY, B. C. 2012. The incorporation of the A2 protein to produce novel Qbeta virus-like particles using cell-free protein synthesis. *Biotechnol Prog,* 28**,** 549-55.

SNYDER, R. D. & EDWARDS, M. L. 1991. Effects of polyamine analogs on the extent and fidelity of in vitro polypeptide synthesis. *Biochem Biophys Res Commun,* 176**,** 1383-92.

SOKOLOVA, E., SPRUIJT, E., HANSEN, M. M., DUBUC, E., GROEN, J., CHOKKALINGAM, V., PIRUSKA, A., HEUS, H. A. & HUCK, W. T. 2013. Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proc Natl Acad Sci U S A,* 110**,** 11692-7.

SON, J. M., AHN, J. H., HWANG, M. Y., PARK, C. G., CHOI, C. Y. & KIM, D. M. 2006. Enhancing the efficiency of cell-free protein synthesis through the polymerase-chain-reaction-based addition of a translation enhancer sequence and the in situ removal of the extra amino acid residues. *Anal Biochem,* 351**,** 187-92.

SONENBERG, N. & HINNEBUSCH, A. G. 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell,* 136**,** 731-45.

SPIRIN, A. S. 2004. High-throughput cell-free systems for synthesis of functionally active proteins. *Trends Biotechnol,* 22**,** 538-45.

SPIRIN, A. S., BARANOV, V. I., RYABOVA, L. A., OVODOV, S. Y. & ALAKHOV, Y. B. 1988. A continuous cell-free translation system capable of producing polypeptides in high yield. *Science,* 242**,** 1162-4.

STANO, P. & LUISI, P. L. 2013. Semi-synthetic minimal cells: origin and recent developments. *Curr Opin Biotechnol,* 24**,** 633-8.

STAPLETON, J. A. & SWARTZ, J. R. 2010. Development of an in vitro compartmentalization screen for high-throughput directed evolution of [FeFe] hydrogenases. *PLoS One,* 5**,** e15275.

STOENESCU, R., GRAFF, A. & MEIER, W. 2004. Asymmetric ABC-triblock copolymer membranes induce a directed insertion of membrane proteins. *Macromol Biosci,* 4**,** 930-5.

STUDIER, F. W. 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif,* 41**,** 207-34.

SUBRAMANIAN, M., TAM, H., ZHENG, H. & TRACY, T. S. 2010. CYP2C9-CYP3A4 protein-protein interactions: role of the hydrophobic N terminus. *Drug Metab Dispos,* 38**,** 1003-9.

SUN, Z. Z., YEUNG, E., HAYES, C. A., NOIREAUX, V. & MURRAY, R. M. 2014. Linear DNA for rapid prototyping of synthetic biological circuits in an Escherichia coli based TX-TL cell-free system. *ACS Synth Biol,* 3**,** 387-97.

SUZUKI, T., EZURE, T., ANDO, E., NISHIMURA, O., UTSUMI, T. & TSUNASAWA, S. 2010. Preparation of ubiquitin-conjugated proteins using an insect cell-free protein synthesis system. *J Biotechnol,* 145**,** 73-8.

SUZUKI, T., ITO, M., EZURE, T., KOBAYASHI, S., SHIKATA, M., TANIMIZU, K. & NISHIMURA, O. 2006. Performance of expression vector, pTD1, in insect cell-free translation system. *J Biosci Bioeng,* 102**,** 69-71.

SZCZESNA-SKORUPA, E., STRAUB, P. & KEMPER, B. 1993. Deletion of a conserved tetrapeptide, PPGP, in P450 2C2 results in loss of enzymatic activity without a change in its cellular location. *Arch Biochem Biophys,* 304**,** 170-5.

SZOSTAK, J. W., BARTEL, D. P. & LUISI, P. L. 2001. Synthesizing life. *Nature,* 409**,** 387-90.

TAKAI, K. & ENDO, Y. 2010. The cell-free protein synthesis system from wheat germ. *Methods Mol Biol,* 607**,** 23-30.

TAN, C., SAURABH, S., BRUCHEZ, M. P., SCHWARTZ, R. & LEDUC, P. 2013. Molecular crowding shapes gene expression in synthetic cellular nanosystems. *Nat Nanotechnol,* 8**,** 602-8.

TARUI, H., IMANISHI, S. & HARA, T. 2000. A novel cell-free translation/glycosylation system prepared from insect cells. *J Biosci Bioeng,* 90**,** 508-14.

TARUI, H., MURATA, M., TANI, I., IMANISHI, S., NISHIKAWA, S. & HARA, T. 2001. Establishment and characterization of cell-free translation/glycosylation in insect cell (Spodoptera frugiperda 21) extract prepared with high pressure treatment. *Appl Microbiol Biotechnol,* 55**,** 446-53.

TAWFIK, D. S. & GRIFFITHS, A. D. 1998. Man-made cell-like compartments for molecular evolution. *Nat Biotechnol,* 16**,** 652-6.

TEE, K. L. & WONG, T. S. 2013. Polishing the craft of genetic diversity creation in directed evolution. *Biotechnol Adv,* 31**,** 1707-21.

TERENIN, I. M., DMITRIEV, S. E., ANDREEV, D. E., ROYALL, E., BELSHAM, G. J., ROBERTS, L. O. & SHATSKY, I. N. 2005. A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. *Mol Cell Biol,* 25**,** 7879-88.

TERWILLIGER, T. C. & EISENBERG, D. 1982. The structure of melittin. II. Interpretation of the structure. *J Biol Chem,* 257**,** 6016-22.

THUMMEL, K. E. 2007. Gut instincts: CYP3A4 and intestinal drug metabolism. *J Clin Invest,* 117**,** 3173-6.

TORIMOTO, N., ISHII, I., TOYAMA, K., HATA, M., TANAKA, K., SHIMOMURA, H., NAKAMURA, H., ARIYOSHI, N., OHMORI, S. & KITADA, M. 2007. Helices F-G are important for the substrate specificities of CYP3A7. *Drug Metab Dispos,* 35**,** 484-92.

VERGERES, G. & WASKELL, L. 1995. Cytochrome b5, its functions, structure and membrane topology. *Biochimie,* 77**,** 604-20.

VIJAYAN, K., DISCHER, D. E., LAL, J., JANMEY, P. & GOULIAN, M. 2005. Interactions of membrane-active peptides with thick, neutral, nonzwitterionic bilayers. *J Phys Chem B,* 109**,** 14356-64.

VON WACHENFELDT, C., RICHARDSON, T. H., COSME, J. & JOHNSON, E. F. 1997. Microsomal P450 2C3 is expressed as a soluble dimer in Escherichia coli following modification of its N-terminus. *Arch Biochem Biophys,* 339**,** 107-14.

WACKER, M., LINTON, D., HITCHEN, P. G., NITA-LAZAR, M., HASLAM, S. M., NORTH, S. J., PANICO, M., MORRIS, H. R., DELL, A., WREN, B. W. & AEBI, M. 2002. N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli. *Science,* 298**,** 1790-3.

WALTER, P. & BLOBEL, G. 1983. Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol,* 96**,** 84-93.

WANG, H. L., CHUNG, T. S., TONG, Y. W., JEYASEELAN, K., ARMUGAM, A., DUONG, H. H. P., FU, F. J., SEAH, H., YANG, J. & HONG, M. H. 2013. Mechanically robust and highly permeable AquaporinZ biomimetic membranes. *Journal of Membrane Science,* 434**,** 130-136.

WANG, R. W., NEWTON, D. J., LIU, N., ATKINS, W. M. & LU, A. Y. 2000. Human cytochrome P-450 3A4: in vitro drug-drug interaction patterns are substrate-dependent. *Drug Metab Dispos,* 28**,** 360-6.

WANG, X., LIU, J., ZHENG, Y., LI, J., WANG, H., ZHOU, Y., QI, M., YU, H., TANG, W. & ZHAO, W. M. 2008. An optimized yeast cell-free system: sufficient for translation of human papillomavirus 58 L1 mRNA and assembly of virus-like particles. *J Biosci Bioeng,* 106**,** 8-15.

WEITZ, M. & SIMMEL, F. C. 2012. Synthetic in vitro transcription circuits. *Transcription,* 3**,** 87-91.

WILLIAMS, P. A., COSME, J., SRIDHAR, V., JOHNSON, E. F. & MCREE, D. E. 2000. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol Cell,* 5**,** 121-31.

WILLIAMS, P. A., COSME, J., VINKOVIC, D. M., WARD, A., ANGOVE, H. C., DAY, P. J., VONRHEIN, C., TICKLE, I. J. & JHOTI, H. 2004. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science,* 305**,** 683-6.

WOOLAWAY, K. E., LAZARIDIS, K., BELSHAM, G. J., CARTER, M. J. & ROBERTS, L. O. 2001. The 5 ' untranslated region of Rhopalosiphum padi virus contains an internal ribosome entry site which functions efficiently in mammalian, plant, and insect translation systems. *Journal of Virology,* 75**,** 10244-10249.

WUU, J. J. & SWARTZ, J. R. 2008. High yield cell-free production of integral membrane proteins without refolding or detergents. *Biochimica Et Biophysica Acta-Biomembranes,* 1778**,** 1237-1250.

XIAO, R. Z., ZENG, Z. W., ZHOU, G. L., WANG, J. J., LI, F. Z. & WANG, A. M. 2010. Recent advances in PEG-PLA block copolymer nanoparticles. *Int J Nanomedicine,* 5**,** 1057-65.

YAMANE, T., IKEDA, Y., NAGASAKA, T. & NAKANO, H. 2005. Enhanced cell-free protein synthesis using a S30 extract from Escherichia coli grown rapidly at 42 degrees C in an amino acid enriched medium. *Biotechnology Progress,* 21**,** 608-613.

YAMAZAKI, H., NAKAMURA, M., KOMATSU, T., OHYAMA, K., HATANAKA, N., ASAHI, S., SHIMADA, N., GUENGERICH, F. P., SHIMADA, T., NAKAJIMA, M. & YOKOI, T. 2002. Roles of NADPH-P450 reductase and apo- and holo-cytochrome b5 on xenobiotic oxidations catalyzed by 12 recombinant human cytochrome P450s expressed in membranes of Escherichia coli. *Protein Expr Purif,* 24**,** 329-37.

YAMAZAKI, H., NAKANO, M., IMAI, Y., UENG, Y. F., GUENGERICH, F. P. & SHIMADA, T. 1996. Roles of cytochrome b5 in the oxidation of testosterone and nifedipine by recombinant cytochrome P450 3A4 and by human liver microsomes. *Arch Biochem Biophys,* 325**,** 174-82.

YAMAZAKI, H., SHIMADA, T., MARTIN, M. V. & GUENGERICH, F. P. 2001. Stimulation of cytochrome P450 reactions by apo-cytochrome b5: evidence against transfer of heme from cytochrome P450 3A4 to apo-cytochrome b5 or heme oxygenase. *J Biol Chem,* 276**,** 30885-91.

YAN, L., HIGBEE, E., TSOURKAS, A. & CHENG, Z. 2015. A simple method for the synthesis of porous polymeric vesicles and their application as MR contrast agents. *J Mater Chem B Mater Biol Med,* 3**,** 9277-9284.

YANG, J. P., CIRICO, T., KATZEN, F., PETERSON, T. C. & KUDLICKI, W. 2011. Cell-free synthesis of a functional G protein-coupled receptor complexed with nanometer scale bilayer discs. *BMC Biotechnol,* 11**,** 57.

YANG, Z., MA, R. R., HUANG, L., ZHU, X. C., SHENG, J. Y., CAI, J., HU, W. & XU, Z. N. 2012. High-level production of soluble adenine nucleotide translocator from Schistosoma japonicum in E. coli cell-free system. *Process Biochem,* 47**,** 395-400.

YIN, G. & SWARTZ, J. R. 2004. Enhancing multiple disulfide bonded protein folding in a cell-free system. *Biotechnol Bioeng,* 86**,** 188-95.

ZAPF, T., TAN, C. W., REINELT, T., HUBER, C., SHAOHUA, D., GEIFMAN-SHOCHAT, S., PAULSEN, H. & SINNER, E. K. 2015. Synthesis and Functional Reconstitution of Light-Harvesting Complex II into Polymeric Membrane Architectures. *Angew Chem Int Ed Engl,* 54**,** 14664-8.

ZELTINS, A. 2013. Construction and characterization of virus-like particles: a review. *Mol Biotechnol,* 53**,** 92-107.

ZEYENKO, V. V., RYABOVA, L. A., GALLIE, D. R. & SPIRIN, A. S. 1994. Enhancing effect of the 3'-untranslated region of tobacco mosaic virus RNA on protein synthesis in vitro. *FEBS Lett,* 354**,** 271-3.

ZHANG, W. H., OTTING, G. & JACKSON, C. J. 2013. Protein engineering with unnatural amino acids. *Curr Opin Struct Biol,* 23**,** 581-7.

ZHANG, Y. & INOUYE, M. 2009. The inhibitory mechanism of protein synthesis by YoeB, an Escherichia coli toxin. *J Biol Chem,* 284**,** 6627-38.

ZHOU, S. F. 2008. Drugs behave as substrates, inhibitors and inducers of human cytochrome P450 3A4. *Curr Drug Metab,* 9**,** 310-22.