

“*Escherichia coli*-based Cell-Free Protein Synthesis of Self-Assembling Particles  
for Vaccine Production and Gene Therapy”

Noelle Angelica Colant

Biochemical Engineering MPhil/PhD

Principal Supervisor: Professor Daniel G. Bracewell

Secondary Supervisor: Dr Stefanie Frank

UCL

Department of Biochemical Engineering

October 2020

**Declaration**

I, Noelle Angelica Colant, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## **Abstract**

The traditional “one-size-fits-all” mass production model commonly used in biologics manufacturing is insufficient to accommodate the advent of personalised medicines and the necessity of on-demand production. The design and validation of novel manufacturing platforms is necessary for on-demand and personalised medicines production.

To address this, an *E. coli*-based cell-free protein synthesis (CFPS) manufacturing platform was developed and applied to self-assembling particles for vaccine and gene therapy production. This in-house CFPS system consistently produces over 400 µg/mL superfolder green fluorescent protein (sfGFP) in 4 hours. A three-step process development strategy that can be completed in under 48 hours was designed and then validated with two products. Using this strategy, sfGFP production was improved by 38% and hepatitis B core antigen (HBcAg) production by 190%.

The CFPS system was then used to produce self-assembling products and iterate upon their construct design. Two tandem-core HBcAg virus-like particles (VLPs), called VLP3 and VLP1, that have been modified to display influenza antigens as universal influenza vaccine candidates were produced and assembled. Using a minimal plasmid backbone designed for CFPS improved titres by 1.8 times over the original VLP1 construct and 1.4 times over the original VLP3 construct. Titres were further increased to over 100 µg/mL for VLP3 when the linkers around the influenza inserts were shortened, although improvements in particle quality were not seen. Further, any constructs with the C-terminal arginine-rich region removed resulted in asymmetric particles of poor quality. Additionally, the three capsid proteins of the adeno-associated virus were produced, which have been shown to form particles *in vitro* and can be used for the delivery of genetic material,

potentially as a gene therapy treatment. Taken together this shows the potential for CFPS systems in the on-demand manufacture of self-assembling vaccine and gene therapy products.

## Impact Statement

Producing self-assembling particles via *E. coli*-based cell-free protein synthesis has led to:

1. A novel process development strategy for cell-free protein synthesis platforms that can be completed in as little as 48 hours. This strategy is based on a robust understanding of the effects of myriad process parameters and how they might be manipulated to rapidly improve titres in any cell-free protein synthesis platform.
2. An increased understanding of the importance of plasmid design and protein design for complex products expressed in cell-free protein synthesis platforms, like tandem-core hepatitis B core antigen virus-like particles. Improvements in plasmid and protein design will lead to the expression of products with fewer lower molecular weight product related impurities and increased titres.
3. A novel synthesis method for tandem-core hepatitis B core antigen virus-like particles and adeno-associated virus serotype 2 virus-like particles allowing for quick and easily scalable deployment of newly designed vaccines and gene therapy products. This would be particularly useful for on-demand production of similarly complex products for personalised medicines or situations where rapid manufacturing is vital like disaster relief, pandemic response, or battlefield delivery.

## **Acknowledgement**

I would like to thank my supervisors, Professor Daniel Bracewell and Dr Stefanie Frank, for sharing their expertise with me, challenging me with new ideas, and supporting me throughout this process. They have made this an exciting and inspiring experience.

I'm grateful for the financial support granted to me by the UCL Overseas Research Scholarship and the UCL Graduate Research Scholarship that allowed me to pursue this incredible opportunity.

I would also like to thank Professor William Rosenberg, Dr Michael Sulu, Dr Beatrice Melinek, Dr Maria Livanos, Dr Stephen Goldrick, Dr Stephen Morris, Dr Lourdes Velez Suberbie, Mark Turmaine, Ian Brown, Christopher Lennon, Ekta Rayani, Alexander van de Steen, Haonan Xu, the UCL iGEM 2018 team, the UCL iGEM 2019 team, and Pilot Plant Week 2019 Team D for not only sharing their scientific knowledge with me but helping me to learn about the joys of being both a mentor and a mentee.

Finally, I'm thankful for all of the encouragement and care my family, especially my parents and my brother, and my friends have provided over the past three years. I could not possibly have done this without their support.

**Table of Contents**

Declaration .....	2
Abstract .....	3
Impact Statement .....	5
Acknowledgement .....	6
Table of Contents .....	7
Index of Figures.....	11
Index of Tables.....	14
Index of Abbreviations .....	15
List of Publications.....	20
1. Introduction and Background.....	22
1.1 Developing an On-Demand CFPS Manufacturing System.....	26
1.1.1 Introduction to Cell-Free Protein Synthesis .....	26
1.1.2 Advantages of CFPS .....	35
1.1.3 Disadvantages of CFPS .....	36
1.2 Designing a Process Development Strategy for CFPS system.....	38
1.3 Using the CFPS system to produce Virus-Like Particles .....	41
1.3.1 Introduction to VLPs .....	41
1.3.2 VLPs as Vaccine Products .....	43
1.3.3 VLPs as Gene Therapy Products .....	52
1.3.4 CFPS of VLPs .....	57
1.4 Aims and Objectives .....	59
2. Materials and Methods .....	61
2.1 Chemicals and Media.....	61
2.2 E. coli cell extract strains and cell bank preparations.....	61
2.3 Plasmid Design and Preparation.....	61
2.3.1 GFP Plasmids .....	61

2.3.2 HBcAg VLP Plasmids .....	62
2.3.3 AAV2 Capsid Protein Plasmids .....	63
2.3.4 Preparation of Plasmids for Cell-Free Protein Synthesis Reactions...	63
2.4 CFPS Reactions .....	64
2.4.1 Cell-Free Extract Preparation .....	64
2.4.2 Cell-Free Concentrated Reaction Mixture Preparation .....	67
2.4.3 Cell-Free Protein Synthesis Reaction Ratios.....	69
2.4.4 Adjusting Process Parameters .....	69
2.4.5 Expressway™ Cell Free Expression System.....	70
2.4.6 Vessels and Scales for CFPS Reactions.....	71
2.5 Analytical Methods .....	72
2.5.1 Fluorescence Analysis to Determine GFP titre .....	72
2.5.2 Total Protein Analysis with Bradford Assay .....	73
2.5.3 SDS-PAGE Analysis.....	73
2.5.4 Dot Blot Analysis .....	74
2.5.5 Western Blot Analysis.....	75
2.5.6 Densitometry Analysis .....	77
2.5.7 ELISA .....	77
2.5.8 Ammonium Sulphate Precipitation of VLPs .....	78
2.5.9 Transmission Electron Microscopy.....	78
2.5.10 Nickel-Nitrilotriacetic Acid Resin Binding for Histidine-Tagged Proteins .....	79
2.5.11 Multivariate Data Analysis .....	81
2.5.12 Design of Experiments for Process Parameter Analysis .....	82
3. Developing a CFPS Manufacturing System for On-Demand Production .....	83
3.1 Introduction .....	83
3.2 Comparing Commercially Available Strains for Cell Extract.....	85
3.3 Evaluating Complex and Minimal Concentrated Reaction Mixtures .....	88

3.4 Examining Plasmids for In Vivo Expression and Plasmids Optimised for CFPS .....	90
3.5 Comparison to Commercial E. coli-based CFPS Kit .....	92
3.6 The Effect of IPTG-Induction in Cell Extract Preparation .....	93
3.7 Scale up of CFPS Reactions to 100 mL.....	97
3.8 Examining IPTG-Induced BL21 Star™ (DE3) Cell Extract Batch-to-Batch Consistency .....	100
3.9 Conclusion .....	101
4. Designing a Process Development Strategy to Improve Titres of CFPS Reactions .....	103
4.1 Introduction .....	103
4.2 Examining a Variety of Process Parameters Using sfGFP.....	106
4.3 Further Analysis Using Hepatitis B Core Antigen.....	109
4.4 Multivariate Data Analysis to Maximise Product Titre .....	114
4.5 Recommended Process Development Method.....	121
4.6 Conclusion .....	122
5. Producing AAV2 VLPs in CFPS .....	124
5.1 Introduction .....	124
5.2 AAV Capsid Protein Production in E. coli-based CFPS .....	125
5.3 Combining Plasmids for Capsid Protein in CFPS Reactions for VLP Formation.....	128
5.4 Purifying Capsid Proteins for In Vitro VLP Formation .....	130
5.5 Conclusion .....	133
6. Characterisation of Product Related Impurities of Tandem-Core Virus-like Particles Produced Using <i>Escherichia coli</i> -based CFPS .....	134
6.1 Introduction .....	134
6.2 Producing Tandem-Core VLPs in an E. coli-based CFPS System .....	136
6.3 Observing Tandem-Core Product Related Impurities .....	140

6.4 Examining the Impact of Plasmid Backbone and Gene Sequence Changes	149
6.5 Conclusion	157
7. Conclusions and Future Work	160
7.1 Review of Project Objectives	160
7.2 Recommendations for Future Work	163
7.2.1 Opportunities for CFPS in Biomanufacturing	165
7.2.2 Addressing Manufacturing Challenges	166
7.2.3 Addressing Business Challenges	174
7.2.4 Addressing Regulatory Challenges	175
8. References	177
9. Appendix A: Protein Sequences and Plasmid Backbones	196
9.1 Hepatitis B Core Antigen Sequence	196
9.2 Tandem-Core K1K1 VLP Sequence	196
9.3 Tandem-Core IAV Vaccine Candidate Sequence and Annotated Modified Constructs	196
9.4 Adeno-Associated Capsid Protein Sequences	203
9.5 Plasmid Backbones	204
10. Appendix B: Using CFPS to Produce “Difficult-to-Express” Products	206
10.1 CFPS for Expression of a Cold Environment Lipase	206
10.2 CFPS of IgG-Specific Endoglycosidase	207
10.3 FujiFilm Diosynth Biotechnologies User Feasibility Study	208

**Index of Figures**

Figure 1.2 Traditional, Stratified, and Personalised Medicines (Hub, 2016) .....	23
Figure 1.3 Cell-Free Protein Synthesis Reaction.....	30
Figure 1.4 Adenosine Triphosphate Regeneration from Different Energy Sources .....	32
Figure 1.5 Cell-Free Protein Synthesis Bioprocess .....	34
Figure 1.6 Tandem-Core Hepatitis B Structure.....	48
Figure 1.7 Influenza A Virus Membrane Proteins .....	50
Figure 1.8 Schematic representation of recombinant protein capsomere constructs adapted from Ramirez et al. 2018.....	51
Figure 1.9 Overview of Triple-Transfection Strategy for Adeno-Associated Virus Production from Vigene Biosciences 2019 .....	54
Figure 1.10 Diagram of Recombinant Adeno-Associated Virus Transduction Pathway from Wang et al, 2019a.....	56
Figure 2.1 Cell Extract Preparation.....	67
Figure 3.1 <i>E. coli</i> Extract Comparison .....	88
Figure 3.2 Concentrated Reaction Mixture Comparison.....	90
Figure 3.3 Plasmid Comparison .....	92
Figure 3.4 Comparison to Expressway™ Mini-Cell Free Expression System.....	93
Figure 3.5 pET Plasmid Expression Under the T7 Promoter with IPTG Induction .....	95
Figure 3.6 IPTG-induced and Non-induced BL21 Star™ (DE3) Extracts with and without Additional T7 RNA Polymerase.....	96
Figure 3.7 Fluorescent Image of SDS-PAGE of CFPS Reactions.....	97
Figure 3.8 Scale up of sfGFP production over three orders of magnitude.....	99
Figure 3.9 Images of sfGFP production.....	99

Figure 3.10 pH and DO levels of bioreactors.....	100
Figure 3.11 BL21 Star™ (DE3) Induced Extract Consistency Test .....	101
Figure 4.1 Process Parameter Effects on Product Titre .....	112
Figure 4.3 Multivariate Data Analysis of Process Parameters .....	118
Figure 4.4 Design of Experiments for sfGFP yields .....	120
Figure 4.5 Design of Experiments for HBcAg yields .....	120
Figure 5.1 CFPS Expression of AAV2 Capsid Proteins.....	127
Figure 5.2 Combining Capsid Protein Plasmids in CFPS Reactions for Capsid Assembly.....	129
Figure 5.3 Ni-NTA Purification of AAV2 Capsid Proteins.....	132
Figure 6.1 CFPS Expression of Tandem-Core HBcAg VLP Constructs .....	137
Figure 6.2 CFPS Reactions to Produce VLP 3 at 30°C and 18°C .....	138
Figure 6.3 CFPS Reactions to Produce VLP 1 at 21°C, 25°C and 30°C .....	138
Figure 6.4 Ammonium Sulphate Precipitation and TEM images of VLP 3.....	141
Figure 6.5 Ammonium Sulphate Precipitation and TEM images of VLP 1 .....	142
Figure 6.8 SDS-PAGE and Western Blot Analysis of Ammonium Sulphate Precipitation of VLP 1 .....	146
Figure 6.9 Lower Molecular Weight Product Related Impurities for Tandem-Core HBcAg VLPs.....	148
Figure 6.10 Ammonium Sulphate Precipitation and TEM images of VLP 3 Variants .....	153
Figure 6.11 Ammonium Sulphate Precipitation and TEM images of VLP 1 Variants .....	154
Figure 6.12 Densitometry and Dot Blot Analysis to determine titre for VLP 3 variants.....	155

Figure 6.13 Densitometry and Dot Blot Analysis to determine titre for VLP 1 variants.....	155
Figure 6.14 Densitometry and Dot Blot Analysis Images of VLP variants .....	156
Figure 7.1 Manufacturing, Business, and Regulatory Challenges for CFPS Production of Personalised Medicines.....	164
Figure 7.2 Challenges Specific to Self-Assembling Particle Production .....	173
Figure 10.1 SDS-PAGE of CFPS Reactions for Cold Environment Lipase .....	207
Figure 10.2 Western blot of ENDO-S2 CFPS Reactions .....	208
Figure 10.3 CFPS Reactions for FDB UFS .....	209
Figure 10.4 Comparing Plasmids Prepared at FDB and UCL .....	211

**Index of Tables**

Table 2.1 Mouse Primary Antibodies for Dot Blots and Western Blots .....75

Table 3.1 *E. coli* strains for Extract Preparation (Doron) (Lucigen, 2018).....86

Table 4.1 Experimental Conditions for Reactions Producing sfGFP Used for MVDA in Figure 4.3A.....119

Table 4.2 Experimental Conditions for Reactions Producing HBcAg Used for MVDA in Figure 4.3B.....119

Table 5.1 Bradford Assay of AAV2 Capsid Proteins Before and After Ni-NTA Purification.....133

Table 6.1 Densitometry and Dot Blot Analysis to determine titre for VLP 3 variants.....157

Table 6.2 Densitometry and Dot Blot Analysis to determine titre for VLP 1 variants.....157

**Index of Abbreviations**

3PGA	3-Phosphoglyceric acid
AAP	Assembly-activating protein
AAV	Adeno-associated virus
AF4	Asymmetric flow field-flow fractionation
AMP	Adenosine monophosphate
ALL	Acute lymphoblastic leukaemia
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
Bio-MOD	Biologically-derived Medicines on Demand
bp	Base pairs
CAR	Chimeric antigen receptor
CCD	Central-composite design
CDMO	Contract development and manufacturing organisations
CECF	Continuous exchange cell-free
CFCF	Continuous flow cell-free
CFPS	Cell-free protein synthesis
CHO	Chinese hamster ovary
CK	Creatine kinase
CMP	Cytidine monophosphate
CoA	Coenzyme A
CP	Creatine phosphate
CPP	Critical process parameter

CQA	Critical quality attribute
CsCl	Caesium chloride
CTP	Cytidine triphosphate
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen
DoE	Design of Experiments
DsbC	Disulphide bond isomerase
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FCC	Face-centred cubic
FDA	Food and Drug Administration
FDB	FujiFilm Diosynth Biotechnologies
FPLC	Fast protein liquid chromatography
FTHM	Future Targeted Healthcare Manufacturing
GMP	Guanosine monophosphate
GFP	Green fluorescent protein
GTP	Guanosine triphosphate
h	Hours
HA	Haemagglutinin
HBcAg	Hepatitis B core antigen

HBsAg	Hepatitis B surface antigen
HEK293	Human embryonic kidney 293
HIC	Hydrophobic interaction chromatography
His	Histidine
HIV	Human immunodeficiency virus
HPV	Human papilloma virus
IAM	Iodoacetamide
IAV	Influenza A virus
IEX	Ion exchange chromatography
InSCyT	Integrated Scalable Cyto-Technology
iP	Inorganic phosphate
IPTG	Isopropyl $\beta$ -D-1-thioglatopyranoside
kDa	Kilo Daltons
K <sub>2</sub> HPO <sub>4</sub>	Potassium phosphate dibasic
KH <sub>2</sub> PO <sub>4</sub>	Potassium phosphate monobasic
KOH	Potassium hydroxide
LAH	Long alpha-helix
LAL	Limulus Amebocyte Lysate
LB	Lysogeny broth media
LPS	Lipopolysaccharide
M2	Matrix 2 protein
MDCK	Madin-Darby canine kidney
mHBcAg	Monomeric Hepatitis B core antigen

MIR	Major insertion region
MLR	Multilinear regression
mRNA	Messenger RNA
MPyV	Murine polyomavirus
MVDA	Multivariate data analysis
NaCl	Sodium chloride
NAD	Nicotinamide adenine dinucleotide
$(\text{NH}_4)_2\text{SO}_4$	Ammonium sulphate
NTA	Nitrilotriacetic acid
OD	Optical density
ORF	Open-reading-frame
PANOx-SP	PEP, Amino acids, NAD, Oxalic acid, Spermidine, and Putrescine
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEP	Phosphoenolpyruvate
pfu	Plaque forming units
PURE	Protein synthesis Using Recombinant Elements
QbD	Quality by Design
RNA	Ribonucleic acid
rpm	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SEC	Size exclusion chromatography
scFv	Single-chain fragment variable
sfGFP	Super folder green fluorescent protein
siRNA	Small interfering RNA
ssDNA	Single-stranded DNA
SMA	Spinal muscular atrophy
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween 20
TBST-M	Tris-buffered saline with Tween 20 in milk
tcHBcAg-GFP	Tandem-core HBcAg VLP with GFP inserts
TEM	Transmission electron microscopy
tRNA	Transfer RNA
UCL	University College London
UFS	User feasibility study
UMP	Uridine monophosphate
UTP	Uridine triphosphate
VLP	Virus-like particle
vRNP	Viral ribonucleoprotein
VP1	AAV2 structural protein 1
VP2	AAV2 structural protein 2
VP3	AAV2 structural protein 3
WT	Wild type
YTPG	Yeast tryptone phosphate glucose medium

## List of Publications

### ***Original Research Papers***

Colant, N, Melinek, B, Teneb, J, Goldrick, S, Rosenberg, W, Frank, S, Bracewell, DG. 2020. A rational approach to improving titre in *E. coli*-based cell-free protein synthesis reactions. *Biotechnology Progress*. (Relates to Chapter 4)

Melinek, B, Colant, N, Stamatis, C, Lennon, C, Farid, SS, Polizzi, K, Carver, M, Bracewell, DG. 2020. Toward a Roadmap for Cell-Free Synthesis in Bioprocessing. *BioProcess International*. (Relates to Chapter 7)

### ***Conference Posters***

Colant, N, Teneb, J, Ogonah, O, Frank, S, Bracewell DG. 2018. Developing an *Escherichia coli*-based Cell-Free Protein Synthesis Platform for Vaccine and Gene Therapy Production. *BioProNET 5th Annual Science Meeting*. London, UK. (Relates to Chapter 3)

Colant, N, Teneb, J, Ogonah, O, Rosenberg, W, Frank, S, Bracewell DG. 2019. Production of Tandem-Core Virus-like Particles Using *Escherichia coli*-based Cell-Free Protein Synthesis. *Biotherapeutics and Vaccines Development Gordon Research Conference*. Galveston, TX, USA. (Relates to Chapter 6)

### ***Conference Presentations***

Colant, N, Teneb, J, Ogonah, O, Frank, S, Ramirez, A, Rosenberg, W, Bracewell DG. 2019. Production of Tandem-Core Virus-like Particles Using *Escherichia coli*-based Cell-Free Protein Synthesis. *ACS BIOT Meeting: Emerging Frontiers in BIOT*. Orlando, FL, USA. (Relates to Chapter 6)

Colant, N, Teneb, J, Melinek, B, Frank, S, Ramirez, A, Rosenberg, W, Bracewell DG. 2020. Rapid Iterative Design of Tandem-Core Virus-Like Particles Using

*Escherichia coli*-based Cell-Free Protein Synthesis. Microbial Engineering II.

Albufeira, Portugal. *Cancelled due to COVID-19*. (Relates to Chapter 6)

Colant, N, Teneb, J, Frank, S, Ramirez, A, Rosenberg, W, Bracewell DG. 2020.

Process development strategy for *E. coli*-based cell-free protein synthesis reactions. American Chemical Society. Virtual. (Relates to Chapter 4)

## 1. Introduction and Background

This doctoral research project was a part of the work supported by the Future Targeted Healthcare Manufacturing (FTHM) Hub. The FTHM Hub is a consortium of universities, vendors, government agencies, industrial biomanufacturers, and innovation centres examining the manufacturing, business, and regulatory challenges that accompany the development and production of new targeted biologic medicines in a quick and cost-effective manner (Figure 1.1). For decades, the biomanufacturing industry has taken a “one-size-fits-all” approach for biologic medicines development. With the advent of new targeted therapies, particularly stratified and personalised medicines, the current approach to drug discovery, process development, and manufacturing will no longer be sufficient (Figure 1.2). New manufacturing strategies will be needed to generate these medicines that will significantly improve patient care and quality of life. That was the goal of this doctoral research project: to establish a novel manufacturing system that could be used to produce self-assembling particles for vaccine production and gene therapy on-demand.

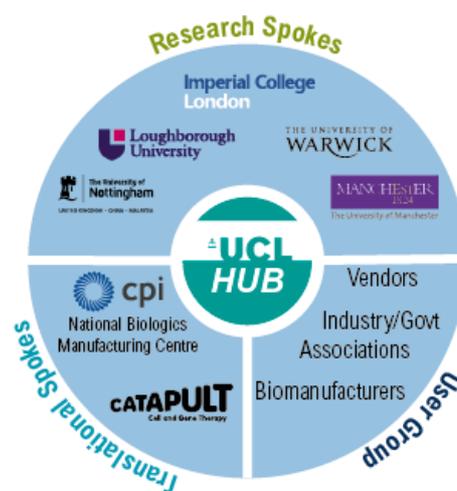


Figure 1.1 The Future Targeted Healthcare Manufacturing Hub Members (Hub, 2016) The Future Targeted Healthcare Manufacturing Hub consists of members involved in three spokes: research, translation (from research to industry) and user groups (industry and/or government). Used with permission from (Hub, 2016).



Figure 1.2 Traditional, Stratified, and Personalised Medicines (Hub, 2016)

Traditional medicines are often manufactured with large patient populations in mind. Stratified medicines are administered to subgroups of patients with similar drug responses and disease profiles while personalised medicines are administered based on an individual patient's disease profile, epigenetic factors, and environmental exposure. Traditional manufacturing methods will need to be adjusted to accommodate more targeted therapies like stratified and personalised medicines. Used with permission from (Hub, 2016).

### *Vaccines and the Need for On-Demand Manufacturing*

According to a report by the World Health Organization (WHO) over 2.5 million lives are saved annually due to vaccines (WHO, 2013). Vaccines are biological agents designed to stimulate an immune response to protect against an infectious disease. They typically consist of a live-attenuated version of the virus, an inactivated version of the virus, a toxoid, or a subunit from the virus – although nucleic acid vaccines may also enter the market soon (Francis, 2018). In the United States alone, there are currently over 80 vaccines licensed by the United States Food and Drug Administration (FDA) and over 3000 ongoing clinical trials investigating various vaccines (FDA, 2020, ClinicalTrials.gov, 2020). As new or more potent infectious diseases emerge, high quality vaccine products will need to be generated quickly and consistently.

### *Gene Therapy and the Need for On-Demand Manufacturing*

Gene therapy is a type of personalised medicine. Personalised medicines are sometimes considered a subset of stratified medicines which are treatments given to subgroups of patients with similar drug responses and disease profiles in order to optimise outcomes. Stratified medicines range from antibodies like Herceptin, which is used to treat HER2-positive breast and gastric cancer patients,

to immune checkpoint inhibitors and cancer vaccines based on genetic analysis (Krzyszczuk et al., 2018). For a personalised medicine, the treatment is engineered specifically for that patient based on their individual disease profile, epigenetic factors, and environmental exposure; it may even include some of their own cells or genetic material.

More specifically, gene therapy is a type of biologic medical treatment in which a gene is either introduced, knocked out, or replaced with a healthy copy inside of a patient. Many of the gene therapy products in development are chimeric antigen receptor T cell (CAR-T) gene therapies. In CAR-T therapies, T cells are removed from the patient's blood stream. They are isolated and inactivated before being engineered with a specific chimeric antigen receptor and allowed to proliferate. Later, they are purified and returned to the patient. The United States FDA has approved two CAR-T gene therapies: Novartis's Kymriah (tisagenlecleucel) and Kite's Yescarta (axicabtagene ciloleucel). Kymriah is a treatment for patients under 25 years of age who have acute lymphoblastic leukaemia (ALL). Yescarta is used to treat non-Hodgkin lymphoma. While these treatments are no doubt life-changing, it can take several weeks to obtain T cells, modify them, and return them to the patient. This makes on-demand manufacturing significantly more challenging. Instead, this project focused on vector-based gene therapies.

In vector-based gene therapies, a vector, usually a virus, is used to transport genetic material directly to the patient's cells. These treatments are currently focused on simple, single-gene defects that result in severely life-limiting conditions for which no viable alternative remedy exists. In December 2017, Spark Therapeutics' Luxturna (voretigene neparvovec-rzyl) was the first vector-based gene therapy to be approved by the United States FDA. It is a drug designed to

treat patients with biallelic *RPE65* mutation-associated retinal dystrophy, a rare form of inherited vision loss that has the potential to lead to complete blindness (FDA, December 19, 2017). Luxturna delivers a normal copy of the *RPE65* gene to retinal pigment epithelial cells using a modified version of a naturally occurring adeno-associated virus (AAV) (FDA, December 19, 2017). Another AAV vector-based gene therapy, Avexis's Zolgensma (onasemnogene abeparvovec-xioi), was approved by the United States FDA in 2019. Zolgensma is used to treat children less than two years old with spinal muscular atrophy (SMA) by delivering a copy of the *SMN1* transgene to boost production of SMN protein, which is deficient in SMA patients (AveXis, 2020). With a ~45% increase in cell and gene therapy clinical trials between 2018 and 2019 and 89% of *in vivo* treatments using AAV, more attempts at FDA approval of AAV gene therapies can be expected in the upcoming years (Catapult, 2019).

#### *Advances in On-Demand Manufacturing*

On-demand production of vaccines and gene therapy is critical in situations like pandemic response, where larger quantities of vaccine products are needed swiftly, or production at a regional pharmacy or hospital bedside to reduce cost and length of hospital stay for gene therapy patients. Advances that would allow for the release of on-demand medicines are just now coming forward. The Integrated Scalable Cyto-Technology (InSCyT) system developed at Massachusetts Institute of Technology uses a benchtop manufacturing scheme to produce over 100 doses of clinical grade drug product in under 100 hours (Crowell et al., 2018). A team at the University of Maryland, Baltimore County invented a portable briefcase-sized system to produce a single dose of drug product in less than 24 hours called Bio-MOD for Biologically-derived Medicines on Demand (Adiga et al., 2018). The University of Utrecht's magistral drug manufacturing strategy significantly

decreases costs for patients who require continual treatment with biologic medicines (Schellekens et al., 2017). Similar systems are also emerging in the industrial space, for example, the NevoLine™ developed by Univercells employs an integrated and automated modular platform for rapid deployment of viral products (Univercells, 2020). In addition to pandemic response and bedside delivery as mentioned above, these systems have been designed to address battlefield delivery where cold chain logistics are difficult, disaster relief where stockpiles are depleted quickly and need to be rebuilt, and production at regional pharmacies to minimise delivery costs. In all cases, these systems decentralise the manufacturing process and decrease the delivery time to the patient. The system developed in this project likewise needed to be rapid and robust, as well as flexible so that it can accommodate a variety of personalised medicines.

## ***1.1 Developing an On-Demand CFPS Manufacturing System***

### ***1.1.1 Introduction to Cell-Free Protein Synthesis***

The on-demand personalised medicines manufacturing system in this project was an *Escherichia coli*-based cell-free protein synthesis (CFPS) platform. In CFPS methods of protein production, living cells are not engaged as part of the expression system. Instead, this *in vitro* production technique engages cell transcription and translation machinery to facilitate recombinant protein expression. CFPS was first used in 1961 to determine the codon for the amino acid phenylalanine (Nirenberg and Matthaei, 1961). Since then it has been employed as a technique to generate a variety of therapeutic proteins, including antibodies, vaccine candidates, and protein biologics [Reviewed in (Ogonah et al., 2017)]. The transcription and translation machinery required in CFPS reactions is derived in one of two ways. Traditionally, it is harvested from cells and formulated into an extract used as one of the raw materials for the CFPS reaction. CFPS extracts

have been generated from a variety of different host organisms including: *E. coli*, wheat germ, rabbit reticulocyte lysate, insect cells (*Spodoptera frugiperda*), *Pichia pastoris*, *Saccharomyces cerevisiae*, Chinese hamster ovary (CHO) cells, HeLa cells, Tobacco, *Neurospora crassa*, *Streptomyces*, *Vibrio natriegens*, *Bacillus subtilis*, *Arabidopsis*, *Pseudomonas Putida*, *Bacillus megaterium*, Archaea, and *Leishmania tarentolae* [Reviewed in (Gregorio et al., 2019)]. In the past decade, another method of CFPS has been developed in which the transcription and translation components have been derived from their Purified Recombinant Elements (PURE) as found in *E. coli* (Shimizu and Ueda, 2010). Although this gives a more predictable and consistent reaction environment because it lacks proteases and nucleases, the PURE system is far more expensive than traditional CFPS using crude extract, and to date, lacks the productivity of the latter (Hong et al., 2014).

CFPS reactions are typically operated in small scale batch reactions in microcentrifuge tubes or micro-well plates. The reactions can also be operated continuously using continuous exchange cell-free (CECF), where the reaction is kept in a central chamber contained by a dialysis membrane surrounded by a reagent-rich solution so that the supply of energy source molecules can be refreshed and by-products, in particular, inorganic phosphate, are able to exit (Stech et al., 2014). Continuous flow cell-free (CFCF), where a substrate rich feed is continuously pumped into the reaction chamber and the products and by-products exit through a semi-permeable membrane, has also been tested (Endo et al., 1992). Additionally, CFPS reagents can be lyophilised which has led to applications such as just-add-water protein production systems, paper-based biosensors, and educational kits (Hunt et al., 2017, Dopp and Reuel, 2018, Thavarajah et al., 2020, Pardee et al., 2014, Pardee et al., 2016, Stark et al., 2018).

Because of the open environment, CFPS reactions are easily manipulated which is beneficial for a variety of applications. Non-standard amino acids can be more readily incorporated into proteins in an open environment (Zimmerman et al., 2014). The lack of cell membranes is advantageous for the expression of proteins that have previously been considered difficult to synthesise like membrane proteins and toxic proteins (Junge et al., 2011, Xu et al., 2005). Translational machinery can be more easily studied; other research groups have experimented with labelling or mutating ribosomes, producing synthetic ribosomes, hybrid ribosome systems, and replacement or removal of tRNAs (Ahn et al., 2006, Yokogawa et al., 2009, Panthu et al., 2018, Jewett et al., 2013). Rapid metabolic pathway engineering for the production of commodity chemicals and natural products has been achieved as well (Jiang et al., 2018).

Although CFPS has numerous applications, there has been relatively little adoption by industry. Several life sciences companies provide pre-made CFPS reaction kits including Arbor Biosciences, Bioneer, Creative Biolabs, Jena Bioscience, LenioBio, New England Biolabs, Promega, Qiagen, and ThermoFisher. The number of organisations currently using CFPS to manufacture their product is noticeably less. Sutro Biopharma is the present industry leader in CFPS of recombinant proteins and monoclonal antibodies. They currently produce three CFPS-manufactured antibody-based products with non-standard amino acids, two of which are in phase 1 clinical trials (Biopharma, 2019). Greenlight Biosciences and EnginZyme each take a slightly different approach to production using CFPS. Greenlight Biosciences uses a cell-free synthesis platform to produce RNA products for biologically derived pesticides for the agricultural industry (Biosciences, 2020). EnginZyme uses enzyme cascades mimicking cellular processes in a packed bed reactor for chemical production (EnginZyme, 2020).

With the onset of the COVID-19 pandemic in the past few months, Swiftscale Biologics has emerged which uses *E. coli*-based CFPS to screen and characterise potential antibody therapeutics (Biologics, 2020).

There are three major components in a CFPS reaction: the cell extract, the concentrated reaction mixture, and the genetic material (Figure 1.3). The cell extract contains important transcription and translation machinery like ribosomes, elongation factors, initiation factors, release factors, etc. *E. coli*-based crude extract CFPS was chosen because it is the most prevalent and well-characterised system present in the research community, and it has been used successfully in industry by Sutro Biopharma. Several different *E. coli* strains have been used to generate CFPS extracts. Some of the most common are modified A19 strains (a derivative of *E. coli* K12) and the BL21 (DE3) strain and its variants, like BL21 Star™ (DE3) and Rosetta™ (DE3); their significance derives from the fact that they are commercially available expression strains and they have been engineered to express T7 RNA polymerase, an important component for transcription in the CFPS reaction (Liu et al., 2005, Kwon and Jewett, 2015, Sitaraman et al., 2004). Other strains of interest include ClearColi® BL21 (DE3), a strain with modifications to the lipopolysaccharide that result in no endotoxin production, and SHuffle T7, which promotes disulphide bond formation (Wilding et al., 2019, Dopp and Reuel, 2019).

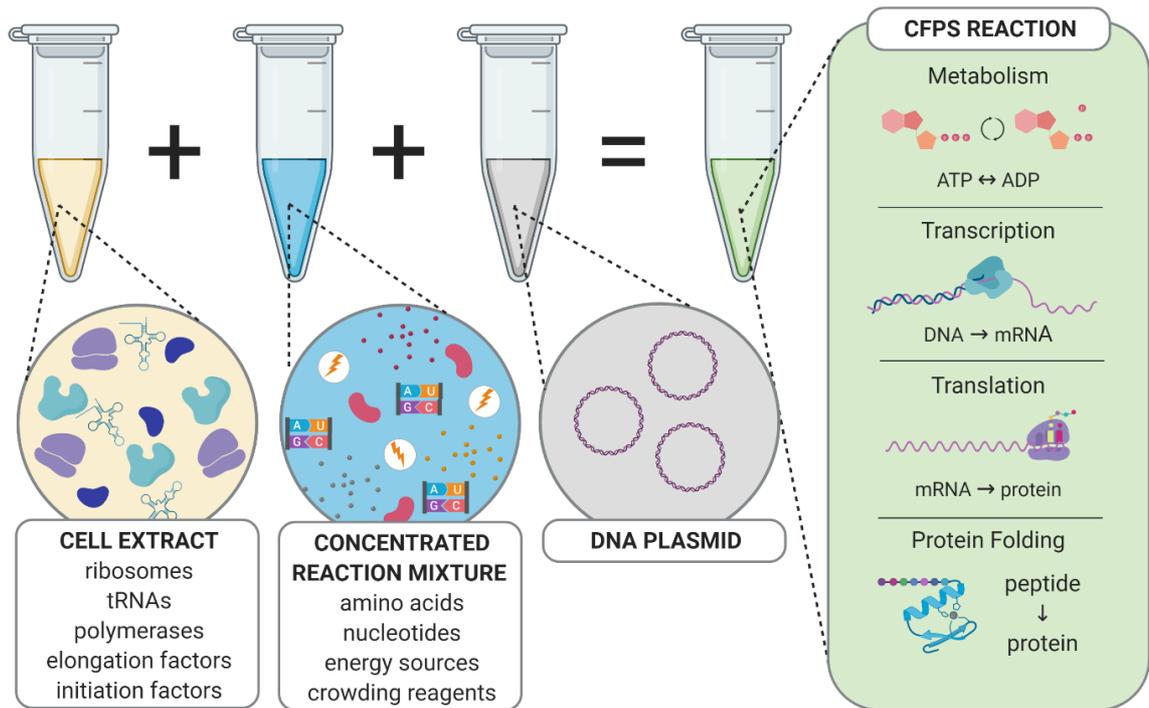


Figure 1.3 Cell-Free Protein Synthesis Reaction

A cell extract containing transcription and translation machinery like ribosomes, tRNAs, polymerases, elongation factors, and initiation factors, a concentrated reaction mixture containing the other necessary components for transcription and translation like amino acids, nucleotides, energy source for ATP regeneration, and crowding reagents, and a DNA plasmid are combined in a CFPS reaction where metabolism, transcription, translation, and protein folding will occur. Created with Biorender.com.

Cell extracts can be prepared in a variety of ways. The following is a general outline of the most common process used for *E. coli* crude lysate CFPS. Cell extracts are prepared from *E. coli* batch cultivations in either shake flasks or fermenters in medium with excess glucose, which results in more productive extracts (Carlson et al., 2014, Jewett and Swartz, 2004, Zawada et al., 2011). Cells are grown to relatively high density, normally with OD<sub>600</sub> measurements in the range of 4 to 6, when the cells are transitioning from growth phase to stationary phase, and then pelleted by centrifugation and frozen. The cells are resuspended in buffer and lysed. A number of different lysis techniques have been successfully utilised for extract production including high pressure homogenisation, French press, sonication, nitrogen cavitation, freeze thaw, and bead beating [Reviewed in

(Gregorio et al., 2019)]. The extract is centrifuged at high speed to remove remaining cell contaminants – usually at 30,000g, from which the name of “S30 extracts” is derived. The extract is aliquoted and flash-frozen with liquid nitrogen or lyophilised. Run-off reactions at 37°C for up to 90 minutes are performed to remove native transcripts and increase extract activity, although some *E. coli* strains, in particular BL21 (DE3) and its derivatives, have shown a decrease in productivity when subjected to a run-off reaction (Krinsky et al., 2016).

Although *E. coli* is usually considered to be a relatively simple host organism, *E. coli* extracts have been manipulated to allow for more complicated products. Disulphide bonds can be formed by pre-incubating extracts with iodoacetamide (IAM) and introducing disulphide bond isomerase (DsbC) into the reaction mixture (Knapp et al., 2006). Recently, glycosylation was achieved in CFPS using glycotransferase and nanodiscs as a scaffold for protein folding (Schoborg et al., 2018).

The next component is the concentrated reaction mixture. The concentrated reaction mixture contains the necessary components for sustaining transcription and translation including: amino acids, nucleotides, energy substrates, salts, T7 RNA polymerase for transcription, additional *E. coli* tRNAs for translation, molecular crowding agents, and, if necessary, chaperones to assist in protein folding (Carlson et al., 2014). Concentrated reaction mixtures are formulated depending on the energy source they use for ATP regeneration (Figure 1.4). Most cell-free protein synthesis platforms use glycolysis to generate ATP where the cofactors coenzyme A (CoA) and nicotinamide adenine dinucleotide (NAD) are supplied along with 3-phosphoglycerate (3PGA), phosphoenolpyruvate (PEP), or pyruvate, although some systems use oxidative phosphorylation with glutamate as the main energy source [Reviewed in (Dopp et al., 2019b)]. Phosphorylated energy

sources result in the accumulation of inorganic phosphate (iP); if the iP concentration in the system exceeds 40-50 mM, the magnesium ion concentration will decrease and protein synthesis will stop (Kim et al., 2006a). Accumulation of iP can be mitigated by adding maltose to the reaction to help with iP recycling (Caschera and Noireaux, 2014). In this project, two of the more common reaction systems, the PANOX-SP System, which relies on glycolysis beginning with PEP, and the Cytomim System, which is intended to mimic the cell cytoplasm and relies on oxidative phosphorylation using glutamate for energy, were tested (Kwon and Jewett, 2015, Cai et al., 2015).

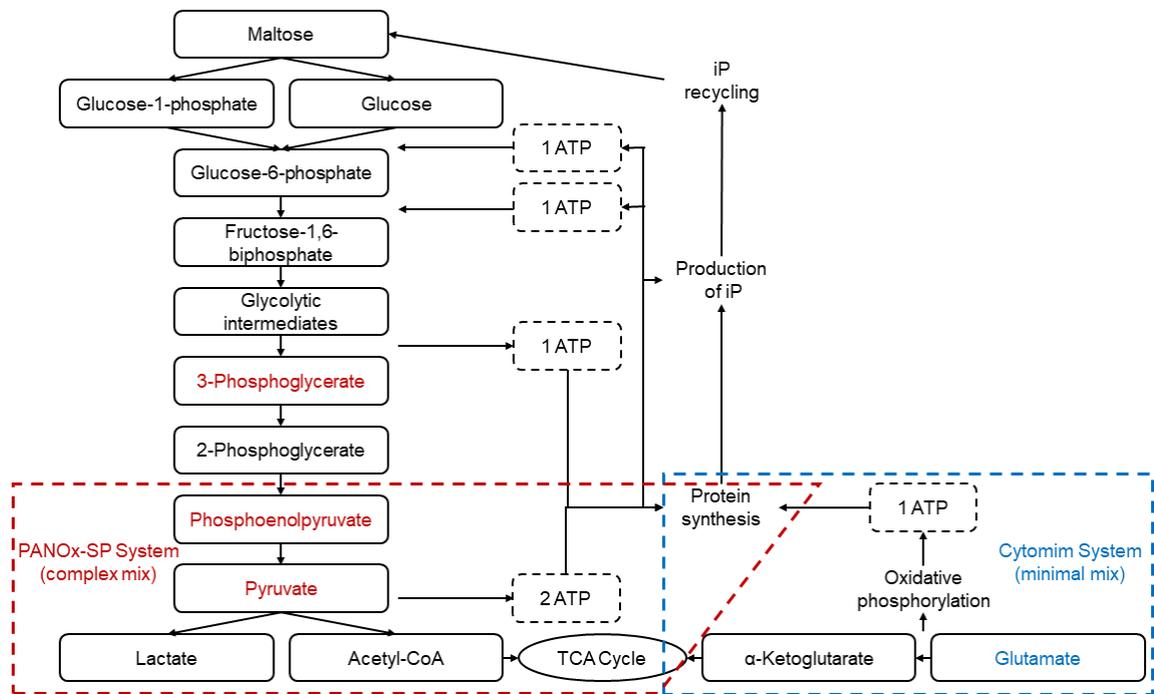


Figure 1.4 Adenosine Triphosphate Regeneration from Different Energy Sources

ATP regeneration in cell-free protein synthesis is supplied by glycolysis using 3-phosphoglycerate, phosphoenolpyruvate, or pyruvate as the main energy source, shown in red. Oxidative phosphorylation can also be used for ATP regeneration where glutamate is the main energy source, shown in blue. The two systems used in this work are shown in the boxes. The PANOX-SP system, a more complex mix, is shown in the box with the red dotted lines. The Cytomim system, a more minimal mix, is shown in the box with the blue dotted lines. This figure has been adapted from (Dopp et al., 2019b, Caschera and Noireaux, 2014, Jewett et al., 2008).

The genetic material is the third and final major component of the reaction mixture. In this project, plasmid DNA was used because it is easy to prepare using

commercially available kits and because plasmids that have been created for *in vivo* cultivations can be transferred to the CFPS system. Just as in *in vivo* processes, plasmid design can impact the titre achievable. While plasmids of any size can be used in CFPS, smaller plasmids have been designed specifically for CFPS reactions that allow high concentrations of plasmid to be added to the reaction (Stark et al., 2018). The preparation of the plasmids is also critical for its performance in a CFPS reaction. Kits from different suppliers use different buffers and may result in plasmid of varying quality (Strychalski and Romantseva, 2020). Many kits use ethanol washes and/or isopropyl alcohol precipitations to purify DNA. If there are still residual amounts of those reagents present when the plasmid is added to the reaction, they may be inhibitory to protein production. Many groups have started using PCR clean up kits alongside their plasmid preparation processes to improve the purity of the plasmid in the reaction.

By developing a better understanding of the three major components in a CFPS system, a better understanding of the CFPS bioprocess as a whole can be achieved (Figure 1.5). This process is the combination of four integrated processes: cell extract preparation, concentrated reaction mixture preparation, plasmid preparation, and the CFPS reaction for protein production. Each individual process is explained greater detail in Chapter 2.

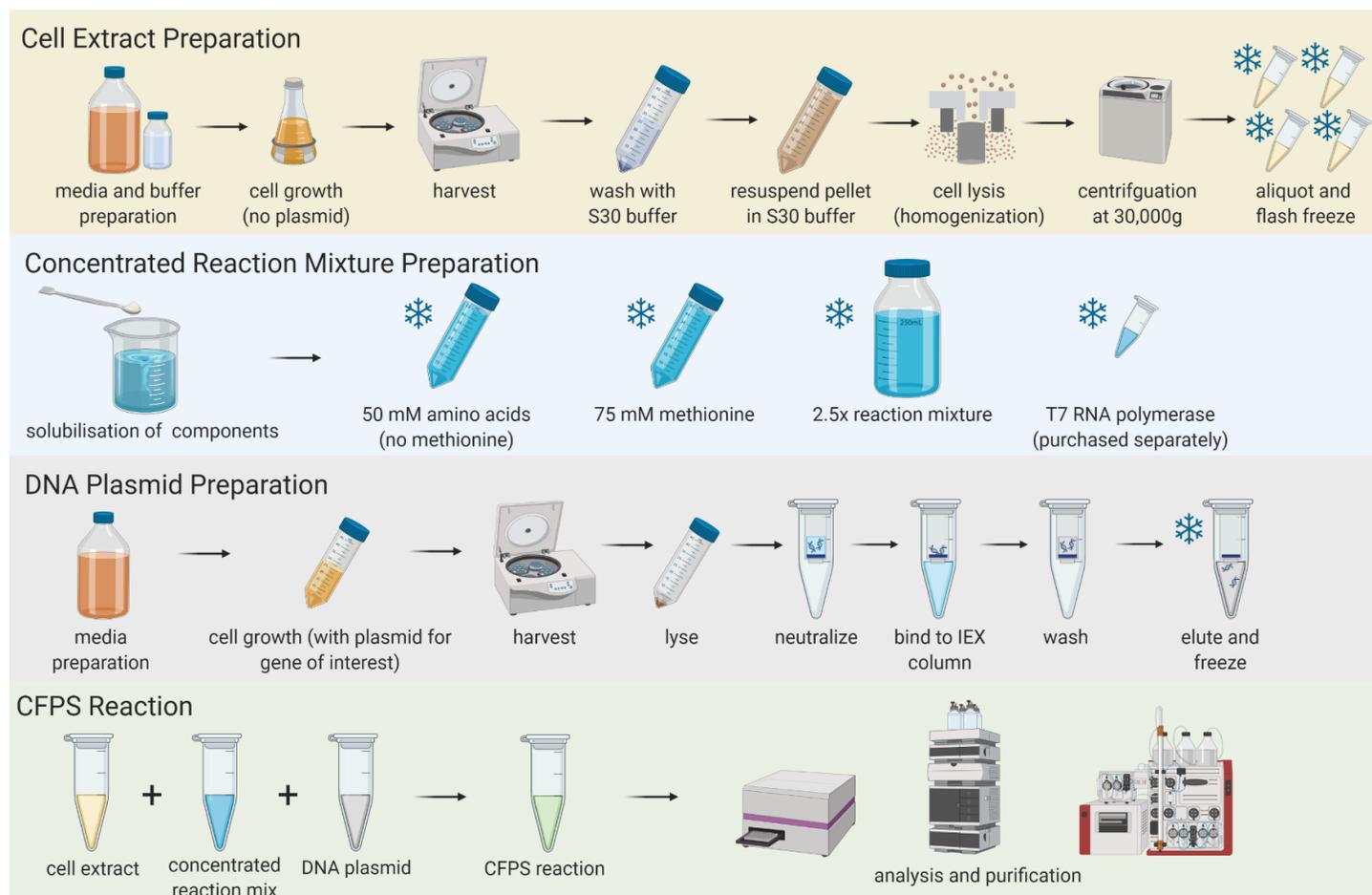


Figure 1.5 Cell-Free Protein Synthesis Bioprocess

The CFPS bioprocess requires cell extract preparation, concentrated reaction mixture preparation (amino acids and T7 RNA polymerase should be prepared separately from the rest of the components), and DNA plasmid preparation. The cell extract and the 2.5x reaction mixture are stored at -80°C. The 50 mM amino acids (no methionine), the 75 mM methionine, the T7 RNA polymerase, and the plasmid DNA are stored at -20°C. These three major components will be combined for the reaction, which can then be analysed and purified. Created with Biorender.com

### 1.1.2 Advantages of CFPS

*E. coli*-based CFPS was chosen as the foundation for the on-demand personalised medicines manufacturing scheme developed in this project because it has three distinct advantages: flexibility, speed, and scalability.

CFPS systems are incredibly flexible. In a traditional cell-based system, plasmid DNA containing the gene of interest would need to be incorporated into the host cell for expression. This would result in a number of clones that would need to be cultivated and compared so that the highest producing clone could be taken forward. In CFPS, there is no cloning necessary. The genetic material is simply added into the reaction. It is not necessary that this genetic material be plasmid DNA. PCR products and mRNA can also be used to express products (Hansen et al., 2016, Schinn et al., 2016). This means that multiple products can be produced at the same time. There is no need to fundamentally change the other reaction components, namely, the cell extract and the concentrated reaction mixture, although some minor modifications may allow for a improvements in titre (this will be examined more closely in Chapter 4).

CFPS reactions are also very quick. Most reactions only take a few hours and still generate relatively high titres because the energy in the system is being used primarily for protein production rather than cell growth and maintenance (Rosenblum and Cooperman, 2014). For example, titres of 1.3 mg/mL chloramphenicol acetyl transferase was produced with an *E. coli*-based CFPS fed-batch system in 2 hours (Kim et al., 2006a). To date, the highest titre achieved in an *E. coli*-based system is 2.3 mg/mL deGFP was produced in batch mode in 10 hours (Caschera and Noireaux, 2014).

CFPS reactions scale linearly provided that the surface area to volume ratio allows for proper oxygen exchange (Voloshin and Swartz, 2005). CFPS reactions

are typically done at small scales, usually less than a few millilitres. However, Sutro Biopharma Inc. was able to produce over 0.7 mg/mL of various model proteins in a 100 L bioreactor in 10 hours while demonstrating consistent scaling starting from 250  $\mu$ L (Zawada et al., 2011).

### 1.1.3 Disadvantages of CFPS

Although CFPS reactions have several advantages, it is also worth recognising the current disadvantages: cost of the system, lack of established downstream processing knowledge, and regulatory concerns. However, improvements are being made in each of these areas.

*E. coli*-based CFPS reactions are currently estimated to cost twice the amount of a CHO cell-based system because they require expensive reagents like the T7 RNA polymerase used for transcription and the phosphorylated energy sources required to sustain the reaction (Stamatis, 2020). In order to cut down on these costs, it has become common practice to produce the T7 RNA polymerase in the lab, rather than purchasing them from a supplier (Dopp and Reuel, 2018). Several *E. coli* strains can also be induced to express T7 RNA polymerase; when these cells that have been induced to express T7 RNA polymerase are formulated into a cell extract, that T7 RNA polymerase remains present in the reaction (Ozawa et al., 2005). Costs can also be decreased by using alternative energy sources like oxidative phosphorylation rather than the phosphorylated energy sources (Cai et al., 2015).

While traditional downstream processing schemes have been used in combination with CFPS reactions, there has been very little work done on optimising these systems for compatibility with CFPS. Many CFPS reactions are performed on a small scale, usually less than a millilitre, where purification would be very difficult. Titres are usually estimated based on fluorescence, if a fluorescent

protein is being produced, or radiolabelled amino acids, where one of the amino acids added to the CFPS reaction, usually leucine or methionine, contains a radioactive isotope of carbon. More complex products, like virus-like particles or bacteriophages, tend to use complex purification schemes involving steps like ultracentrifugation that are not scalable and cannot be incorporated into an on-demand manufacturing scheme (Bundy et al., 2007). However, there is a great opportunity for downstream processing of CFPS reactions. Although intended to mimic the cell, CFPS reactions are significantly more dilute than cell cytoplasm which should allow for a simpler purification strategy (Underwood et al., 2005). It is possible to imagine a versatile system where several products are generated via CFPS using the same extract and reaction mix. With such a similar background, it is conceivable that the downstream process could be the same (or nearly the same) for each product. Some groups are already working towards this goal by designing small scale chromatography columns that would be compatible with low-volume CFPS reactions (Andar et al., 2019).

At the moment, there are no products on the market that are manufactured using CFPS, and therefore there is no regulatory precedent. In setting this precedent, there are a number of concerns that will need to be addressed, particularly with regard to maintaining consistency of the process and the product. A recent study performed by the US Army Combat Capabilities Development Command Chemical Biological Center demonstrated that CFPS reaction titres vary greatly depending on the laboratory and the operator (Cole et al., 2019). However, more automated systems with better environmental controls may be able to mitigate these fluctuations (Georgi et al., 2016, Quast et al., 2015, Caschera et al., 2011). Still, this relies on consistency of the reaction reagents. While establishing detailed standard operating protocols to minimise batch-to-batch variation for the

cell extract, the concentrated reaction mixture, and the genetic material is no doubt a step in the right direction, additional measures of activity or purity may need to be developed in order to adhere to regulatory standards (Dopp et al., 2019a). Hopefully, these areas of concern will be addressed as Sutro Biopharma continues with its clinical trials.

### **1.2 Designing a Process Development Strategy for CFPS system**

To rapidly deploy new biologic medicines on-demand, a robust and reliable process development strategy is needed. Process development strategies are critical for ensuring the consistency and maximal titre output in any biomanufacturing process. The strategy designed in this project was informed by the principles of Quality by Design (QbD). This means identifying critical quality attributes (CQAs) and critical process parameters (CPPs). A CQA is a “physical, chemical, biological, or microbiological property or characteristic that should be within an appropriate limit, range, or distribution to ensure the desired product quality” and a CPP is a “a process parameter whose variability has an impact on a critical quality attribute and therefore should be monitored or controlled to ensure the process produces the desired quality” (ICH, 2009).

Because CFPS systems are just now being used as manufacturing schemes, there has been relatively little work done on devising appropriate process development strategies. In other words, the relevant CPPs for the system and their impact on the corresponding CQAs is largely unknown. In this project, those CPPs and CQAs have begun to be elucidated. The plasmid design, the composition of the concentrated reaction mixture, and the strain of the cell extract were key areas of focus. For simplicity’s sake, the same preparation protocol was followed for all cell extracts that were produced, regardless of the *E. coli* strain being used – as long as the preparation methods were consistent no dramatic

differences in cell extract activity should be expected (Cole et al., 2019). However, other groups suggest optimising magnesium and potassium concentrations for each batch of extract and adjusting the lysis conditions depending on the volume of extract being prepared which may improve yield by up to 30% (Zubay, 1973, Shin and Noireaux, 2010, Kwon and Jewett, 2015). The components in the concentrated reaction mixture were not examined individually as several other groups have already provided guidance in this area. A full history of reaction mixture formulations for *E. coli*-based CFPS reactions with specific focus on more recent attempts to drive down cost and simplify the system can be found in (Dopp et al., 2019b). The impact of the concentration of the plasmid, the amount of cell extract, the pH of the concentrated reactions mixture, the temperature of the reaction, and the length of the reaction on product titre were also observed.

The process development strategy designed in this project is a starting place for risk management in CFPS systems because changes will be made as CPPs and CQAs are better understood. This strategy was designed to be flexible so that new information could be easily incorporated. At the moment, the CQAs and CPPs are largely dependent on the product of interest. If the desired product requires the incorporation of non-standard amino acids, then the concentration of non-standard amino acids in the reactions would be an important process parameter. Likewise, if the product required disulphide bonds, the amount of iodoacetamide (IAM) incubated with the extract, the length of IAM incubation, and the concentration of disulphide bond isomerase (DsbC) in the reaction would all be worthwhile parameters to observe.

A QbD-based approach will also require the use of systematic analyses to better understand the impact of certain CPPs on CQAs. By using tools like multivariate data analysis (MVDA) and design of experiments (DoE), the

interactions among several process parameters cannot only be examined, but the process development timeline may also be shortened, which can reduce the overall time to market. In this project, MVDA was used to examine which process parameters have the most significant impact on titre and then DoE was applied to optimise for the maximum titre.

Multivariate data analysis (MVDA) is an analysis approach in which several input variables and output variables are simultaneously monitored. Using this approach, the input variables that are dependent on other input variables and how those interactions affect output variables can be better understood. For example, MVDA has been used previously to examine cell line and bioreactor operating conditions in order to determine the root cause of trisulphide bond formation in an antibody-peptide fusion product (Goldrick et al., 2017). Because there is a multitude of process parameters that can be examined in a CFPS platform, MVDA is a useful tool for normalising across parameters and determining which parameters have the greatest impact on the output parameter of interest, titre. This allows for the selection of the parameters that are most influential that can then be optimised further.

MVDA goes hand-in-hand with design of experiments (DoE). DoE is a strategy often used in biomanufacturing in which multiple input variables are simultaneously manipulated within a given design space to minimise the number of experiments performed. The design space is constrained by an upper and lower bound for each input variable. An analysis of variance (ANOVA) is completed to identify which input variables have the most significant impact on the output variable and whether there are interactions between the input variables. Several DoE optimisations have been performed with CFPS reactions to improve titres. One group was able to double the titre of wild type green fluorescent protein (GFP)

attained in their system by using a central composite design (CCD) to optimise amino acids and energy substrate concentrations in the concentrated reaction mixture (Pederson et al., 2011). A face-centred cubic (FCC) CCD was used to optimise the induction time and cultivation time for BL21 Star<sup>TM</sup> (DE3) extract preparation so that it could be completed in a single workday (Dopp and Reuel, 2018). Another group used a machine learning algorithm to perform evolutionary DoE and boost eGFP titres by ~350% (Caschera et al., 2011). DoE strategies have also been used previously to optimise IAM incubation and DsbC concentrations for improved disulphide bond formation in therapeutic proteins, antibody fragments, and vaccine fusion proteins (Goerke and Swartz, 2008).

### ***1.3 Using the CFPS system to produce Virus-Like Particles***

#### ***1.3.1 Introduction to VLPs***

Due to their ability to secure, transport, and deliver genetic material, viruses like AAV have long since been of interest to researchers. Their applications range from materials science and nanotechnology to virotherapy and biopharmaceutical development. However, most of the viruses utilised in these technologies need to be non-infectious and should not contain their own DNA or RNA. This has stimulated exploration into the manufacture of virus-like particles (VLPs). VLPs are self-assembling conformational epitopes of viral structural proteins. These particles resemble viruses in many ways, but unlike viruses, they lack pathogenicity as well as replicative abilities because they do not contain their own genetic material (although they may carry non-native nucleic acids). Due to their similarity in structure to their corresponding viruses, VLPs have the ability to invoke strong B cell and T cell responses even when delivered in low doses (Bachmann and Jennings, 2004). This also makes them valuable vaccine candidates because their efficacy could result in a reduction in vaccine costs due to the fact that only a small

amount of material is necessary to generate a strong immune response, even in the absence of adjuvants (Bachmann et al., 1993). Additionally, most VLPs are relatively small, less than 40 nm in diameter, making their uptake into lymphatic vessels and subsequent detection by the immune system more likely compared to larger particles (Oussoren et al., 1997). Existing vaccines on the market that utilise VLPs include Merck's Recombivax HB and GlaxoKlineSmith's Engerix, both of which use hepatitis B surface antigen (HBsAg) VLPs to confer protection against the hepatitis B virus, as well as Merck's Gardasil and GlaxoKlineSmith's Cervarix, which prevent human papilloma virus (HPV) infection using a series of HPV VLPs of the most high-risk strains (Roldão et al., 2010, Ramqvist et al., 2007). Currently, over one hundred VLPs for over forty different viruses, including the human immunodeficiency virus (HIV), the influenza virus, the Ebola virus and the Zika virus, are being explored as potential vaccine candidates [Reviewed in (Zeltins, 2013, Roldão et al., 2010, Boigard et al., 2017)]. They have been produced in a variety of different hosts including bacteria, yeast, insect, plant, and mammalian cells [Reviewed in (Zeltins, 2013)].

However, cell-based production processes can take a substantial amount of time, anywhere from a few days to a few weeks, and even in high producing processes, the VLPs may have an inconsistent architecture and composition (Pattenden et al., 2005). Additional variation can be seen when cell-based processes are transferred to larger scales. These issues may be mitigated by using a CFPS platform which should allow for the generation of higher product titres over a shorter length of time, the production of more consistently assembled particles, and the development of more scalable processes.

VLPs are purified using a number of different techniques. The assembled particles can be precipitated using ammonium sulphate or polyethylene glycol

(PEG) to remove host cell proteins. Ultracentrifugation is typically used at the lab-scale to remove product related impurities from small volume processes. This technique is performed with sucrose gradients or caesium chloride (CsCl) gradients. At larger scales, chromatography is often used. Depending on the product and its properties, size exclusion chromatography (SEC), ion exchange chromatography (IEX), hydrophobic interaction chromatography (HIC), or affinity chromatography may be employed. An additional filtration step may be necessary depending on the purity required. However, in many cases the VLPs are disassembled into their monomeric form, purified (usually via a chromatographic method), and reassembled to ensure consistency.

Following purification, a variety of characterisation techniques are used depending on the properties of the VLPs. Assays like western blots, ELISAs, and Ouchterlony's double radial immune diffusion can be used to demonstrate the presence of the VLPs. Electron microscopy can be used to establish visual proof of the particles. Mass spectrometry can be used to measure molecular mass and identify differences in post-translational modifications. Particle size and size distribution can be determined with dynamic light scattering (DLS) or asymmetric flow field-flow fractionation (AF4) (Chaun et al., 2008).

### 1.3.2 VLPs as Vaccine Products

Not only can VLPs be administered as vaccines to their corresponding virus, but they can be used as scaffolds to display other antigens. These epitopes are either fused genetically to VLP subunit proteins or attached by covalent or noncovalent mechanisms to the VLP surface (Bachmann and Jennings, 2004). Viral coat proteins for well-studied viruses like murine polyomavirus (MPyV), bacteriophage Q $\beta$ , bacteriophage AP205, hepatitis B virus, and HPV have been used to display structural and non-structural proteins from more complex viruses

and cancer-related proteins in order to provoke an immune response against viral infection, to induce a T-cell response to prevent the formation of viral tumours, and to prevent the formation of cancer of non-viral origin by inciting an immune response to a self-antigen [Reviewed in (Ramqvist et al., 2007)]. In one study, nicotine derivatives were covalently coupled to Q $\beta$  VLPs in order to lower nicotine levels in the brain and reduce the smoking habits of patients (Maurer et al., 2005). However, some foreign antigens are too large or too hydrophobic and can cause the structural disintegration of the VLP when attached. This has led to the design of hepatitis B tandem-core VLPs as an attempt at a universal influenza vaccine.

### 1.3.2.1 Traditional Influenza Vaccines

One of the first documents to detail what was most likely influenza was written in the fifth century BC. Since then, several influenza pandemics have swept across the globe. In 1933, the cause of the disease was identified as a virus (Smith et al., 1933). There are four types of influenza: A, B, C, and D (Hause et al., 2014, Palese and Shaw, 2007). In this project, vaccines for the influenza A virus (IAV) were examined. IAV is an enveloped virus in the *Orthomyxoviridae* family that contains a negative sense RNA segmented genome encoding viral genes for 11 proteins: haemagglutinin, neuraminidase, matrix 1, matrix 2, nucleoprotein, non-structural protein 1, non-structural protein 2, polymerase acidic protein, polymerase basic protein 1, polymerase basic protein 2 and polymerase basic protein 1-F2 (Palese and Shaw, 2007). The haemagglutinin, neuraminidase, and matrix 2 proteins make up the lipid bilayer that forms the viral envelope and they are supported by the matrix 1 protein (Samji, 2009). Inside of the matrix 1 protein shell are the eight viral ribonucleoprotein (vRNP) complexes composed of viral RNA wrapped around a nucleoprotein and bound by the viral polymerase which

consists of polymerase acidic protein, polymerase basic protein 1 and polymerase basic protein 2 (Dou et al., 2018).

Currently, influenza vaccines are made by tracking antigenic variation across species, especially humans, pigs, and birds, in the haemagglutinin and neuraminidase proteins found on the influenza virus (Gerdil, 2003). The WHO established an international surveillance network in 1947 that meets biannually to determine the vaccines that will be dispatched in the northern hemisphere and the southern hemisphere (Gerdil, 2003). This allows for a six-month vaccine production cycle. In 2003, an estimated 250 million doses were brought to market in over 100 countries (Gerdil, 2003). In 2019, over 162 million doses were manufactured for use in the United States alone (Prevention, 2019). The vaccine contains two strains of influenza A (H1N1 and H3N2) and at least one strain of influenza B. It is made in embryonated chicken eggs; the virus is grown, harvested, inactivated using formalin or (beta)-propiolactone, purified by ultracentrifugation, and split (Gerdil, 2003). It is then tested in clinical trials in Europe and released.

Although the inactivated virus from the egg-based manufacturing process is the most common vaccine on the market, there are now several other influenza vaccines generated by a range of available production methods. A vaccine using a live attenuated virus, which is often administered by nasal spray, is also manufactured in eggs (Carter and Curran, 2011). In 2012, the United States FDA approved the first cell-based vaccine, Flucelvax (also known as Optaflu) in which viruses are manufactured in Madin-Darby canine kidney (MDCK) cells and then inactivated (Manini et al., 2015). A recombinant influenza vaccine composed entirely of genetically designed haemagglutinin molecules called Flublok has also been approved by the United States FDA (Cox and Hollister, 2009).

All of the vaccines mentioned in the preceding paragraph are based on the haemagglutinin subunit present in the virus, specifically the head of the haemagglutinin globule, often referred to as HA1. Because the HA1 region is so variable, new vaccines are needed each year and there is no guarantee that the vaccine will have high efficacy given that the predicted strain may or may not match the emergent seasonal strain. Attempts to create more potent and less strain dependent vaccines have included DNA vaccines, peptide-based vaccines, vector-based vaccines, and virus-like particle vaccines (Carter et al., 2019, Francis et al., 2015, Hoelscher et al., 2006, Rezaei et al., 2013). The tandem-core VLP-based universal IAV vaccine, which targets less variable regions of IAV, the matrix 2 protein ectodomain (M2e) and the haemagglutinin stalk globule (HA2), was examined in this project.

#### *1.3.2.2 Tandem-Core Influenza Vaccines*

In the tandem-core VLPs developed, a linker protein consisting of a repeated glycine-glycine-serine sequence fuses two hepatitis B core antigen (HBcAg) monomers together, stabilising the dimer and allowing for the insertion of larger or more hydrophobic proteins in the insertion region (Peyret et al., 2015). The HBcAg tandem-core VLP scaffold is a flexible platform that could present any number of antigens for future therapeutics treatments.

The typical hepatitis B virus has a capsid shell composed of surface antigen (HBsAg) with a core antigen lattice underneath (HBcAg) (Figure 1.6A). HBcAg VLPs were the first VLPs to be visualised underneath a microscope (Cohen and Richmond, 1982). Since their discovery, they have been used as a model VLP in several studies. The virus appears in eight alphabetically labelled genotypes of different nucleotide sequence homology and four serotypes based on the type of epitopes found in the envelope protein. Serotypes are determined by group-

specific antigenic determinant 'a' and two pairs of mutually exclusive subtype specific determinants, 'd/y' and 'w/r'. Certain serotypes only appear in certain genotypes; *ayw* is present in all genotypes except 'C', *adw* occurs in all genotypes except 'D' and 'E', and serotypes *ayr* and *adr* can only be found in genotype 'C' (Acharya and Batra, 2005). Serotypes *adw* and *ayw* were both used in this project.

In the HBcAg VLP, 90 dimers or 120 dimers (composed from individual HBcAg monomers) self-assemble into a T=3 or a T=4 icosahedral capsid shell through trimeric and pentameric intermediates (Holmes et al., 2015, Ludgate et al., 2016). These shells are quasi-equivalent and demonstrate a very small difference in size, 31.8 nm in diameter for the T=3 shell and 34.8 nm in diameter for the T=4 shell (Zlotnick et al., 1999). In HBcAg tandem-core VLPs, two HBcAg monomers form a dimer and are linked by a poly-glycine-serine linker protein (Figure 1.6B) (Ramirez et al., 2018). All of the VLPs used in this project contain a linker with seven repeats of glycine-glycine-serine sequence, although constructs with different numbers of repeats exist. The linker stabilises the dimer and assists in assembling other dimer units into a VLP (Ramirez et al., 2018). The VLPs can be formed of homo-tandem dimers or hetero-tandem dimers. The homo-tandem dimers are composed of two C-terminally truncated molecules terminating at amino acid 149 while the hetero-tandem dimers have one truncated molecule and one full length molecule with an arginine-rich nucleic acid binding sequence extending to amino acid 185, which affects particle assembly and stability (Figure 1.6C and Figure 1.6D) (Peyret et al., 2015). Generally, the hetero-tandem dimer constructs result in more homogenous particles than the homo-tandem dimer constructs (Peyret et al., 2015). In each dimer, there are two major insertion regions (MIRs), where antigens from various other viruses can be bound. Because the linker

protein gives the particle more stability, larger and more hydrophobic antigens can be incorporated into the VLP.

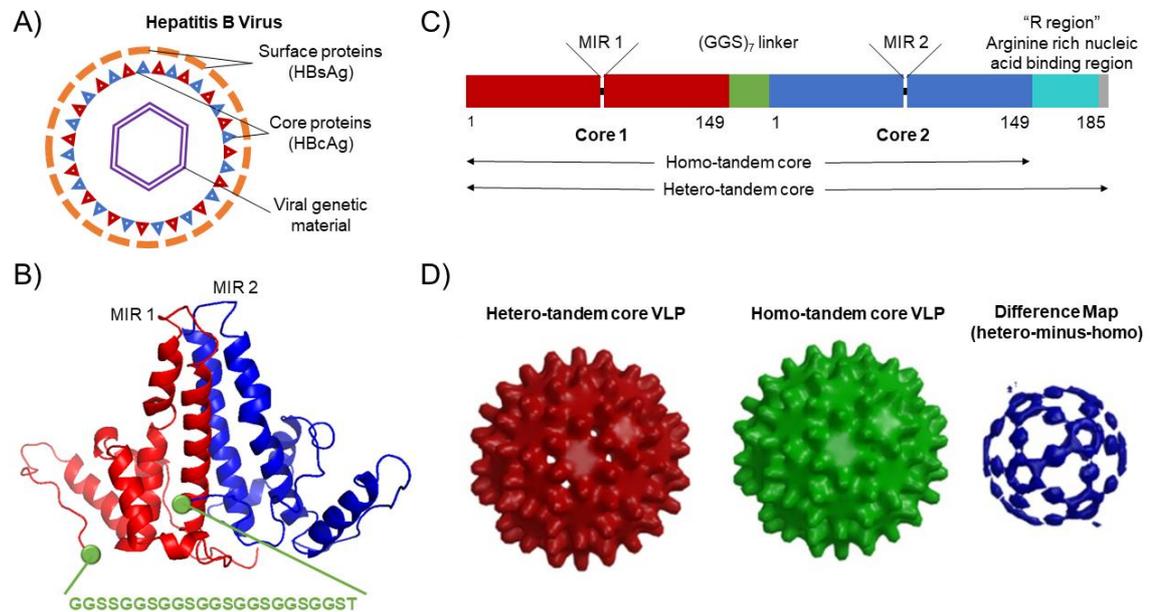


Figure 1.6 Tandem-Core Hepatitis B Structure

(A) Hepatitis B virus with surface antigen (HBsAg) in orange, core protein (HBcAg) in red and blue, and viral DNA in purple (B) HBcAg tandem-core dimer with glycine-glycine-serine linker shown in green and MIRs labelled (C) Difference between homo-tandem core and hetero-tandem core modified from Peyret et al. (D) Surface rendered views of the hetero-tandem core VLP (red), the homo-tandem core VLP (green), and the difference between the two (hetero-minus-homo) (blue) modified from Peyret et al. The VLPs are ~35 nm in diameter. Used with permission from (Peyret, et al., 2015).

For the tandem-core universal IAV vaccine, the IAV matrix 2 protein ectodomain (M2e) and haemagglutinin stalk globule (HA2) have been incorporated into the MIRs (Figure 1.7A). The matrix 2 protein is an ion channel protein found in the envelope of IAV and it is relatively consistent across influenza A strains. M2e is the outermost segment of the matrix 2 protein; in nature, these proteins are joined together as a tetramer, but they are display individually in the tandem-core constructs (Figure 1.7C). In previous studies, mice immunised with M2e were protected against the influenza A virus (Mozdzanowska et al., 2003, Neiryneck et al., 1999). However, M2e exhibits low immunogenicity on its own; it is much more

potent when fused with a carrier (Lega et al., 2016). The three M2e proteins used in the MIR of tandem-core VLPs in this project correspond to the H1N1, H5N1, and H11N9 variants (Ramirez et al., 2018).

HA2, occasionally referred to as the long alpha-helix (LAH), is a helical subunit of the haemagglutinin stalk globule that is very similar across all influenza strains. Mice who were infected with a carrier protein presenting the HA2 antigen exhibited resistance to three different influenza strains (Fan et al., 2015). At a 2017 workshop held by the National Institute of Allergy and Infectious Disease, inclusion of the HA2 stalk protein was recommended for a future universal influenza vaccine as that would allow protection across different HA subtypes (Paules et al., 2017). The HA2 stalk proteins are normally joined together as a trimer, but a single monomer is displayed in the tandem-core constructs (Figure 1.7B).

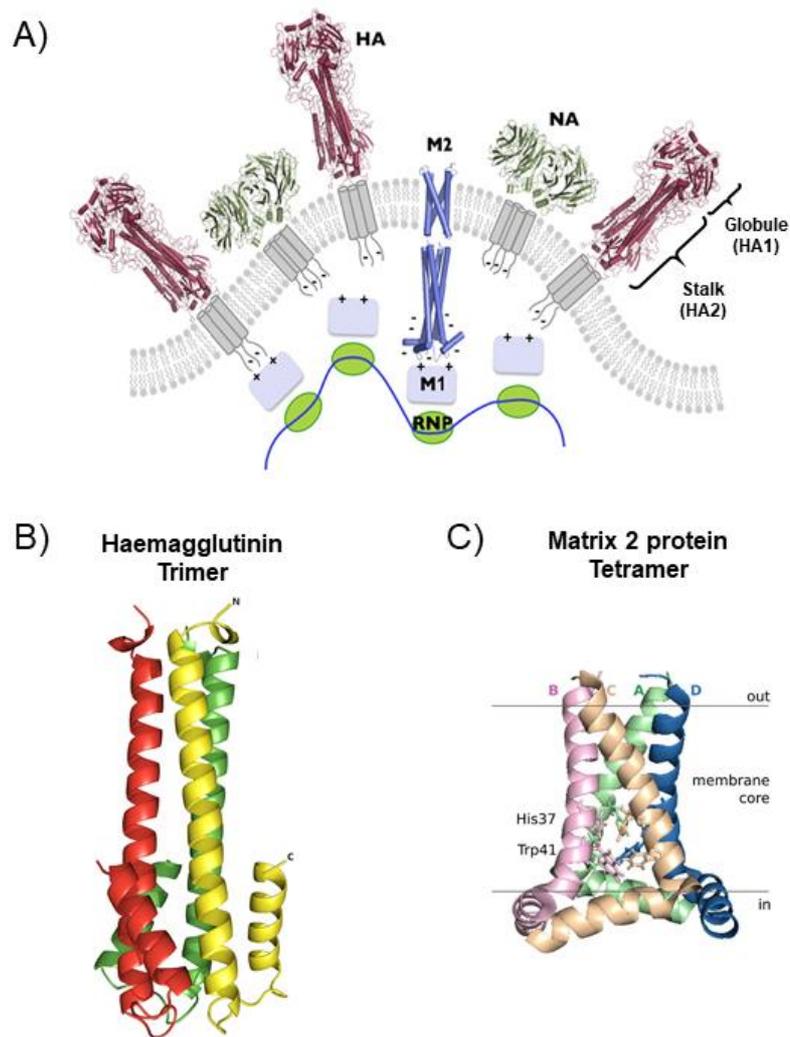


Figure 1.7 Influenza A Virus Membrane Proteins

(A) The influenza A membrane proteins include neuraminidase (NA), matrix protein 2 (M2) and haemagglutinin (HA) which consists of two proteins, the globule (HA1) and the stalk (HA2). These membrane proteins are associated with other viral proteins including matrix protein 1 (M1) and ribonucleoprotein (RNP). This figure has been modified from (Pielak and Chou, 2011). (B) The structure of the haemagglutinin stalk protein trimer. Each individual monomer is shown in a different colour. Used with permission from (Lu et al., 2018). (C) The structure of the matrix protein 2 (M2) tetramer, which forms an ion channel. Each individual monomer is shown in a different colour. Used with permission from (Martinez-Gil and Mingarro, 2015).

There are four tandem-core VLPs that are of particular interest to this project (see Figure 1.8). The first is the K1K1 VLP which contains a single amino acid, lysine, in its MIRs rather than an antigen, making the MIRs essentially “empty”. VLP 3 contains three M2e antigens from different IAV strains in one MIR – the other MIR is empty, it does not contain a lysine. VLP 3 has been successfully produced, purified, and quantified on multiple occasions (Ramirez et al., 2018). Much like VLP 3, VLP 1 contains three distinct M2e antigens in one MIR, but it also

has a single HA2 antigen in the other MIR. This combination of antigens should allow for a more potent vaccine. Unfortunately, because of the number and size of antigens present in VLP 1, it has yet to be produced with the same level of success seen with VLP 3. In a recent study, mice immunised with both VLP 1 and VLP 3 separately had a 100% survival rate when infected with an H1N1 IAV, although the vaccine did not result in sterilising immunity (Ramirez et al., 2018). Finally, a tandem-core VLP with GFP in one of its MIRs (the other is empty) is also of interest as its fluorescent properties allow for a large variety of analytical approaches for titre and assembly determination.

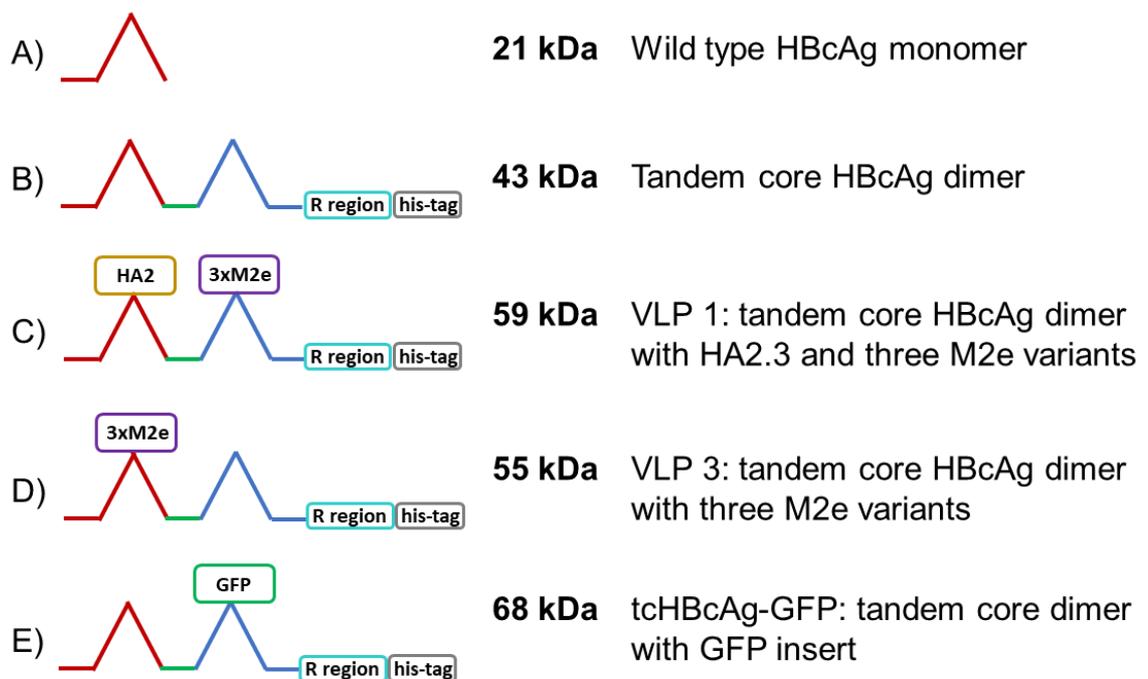


Figure 1.8 Schematic representation of recombinant protein capsomere constructs adapted from Ramirez et al. 2018.

Wild type hepatitis B virus core antigen (HBcAg) (A) self assembles into a virus like particle. Tandem-core (B) is made up of two HBcAg linked together and expressed as a single protein, this construct is also known as the K1K1 variant. VLP 1 (C) incorporates HA-stalk antigen (HA2) from a H1N1 strain in the first major insertion region (MIR) and three variants of M2e antigen in the second MIR. VLP 3 (D) contains three variants of the M2e sequence in one MIR. (E) Tandem-core VLP with a GFP insert in one MIR. Adapted from (Ramirez et al., 2018)

The tandem-core technology is currently licensed to iQur Ltd.; their production scheme utilises a *Pichia pastoris* expression system. Many of their

VLPs have shown assembly issues in *P. pastoris*, forming “lumps” which do not allow for antigens to be displayed properly and leading to less effective VLPs. These assembly issues may be due to steric effects caused by the influenza inserts in the MIRs. Both the hemagglutinin stalk protein and the matrix 2 ectodomain protein are part of multiprotein structures in nature. Those interactions may be occurring to some extent even when these proteins are incorporated into a larger VLP structure which may lead to aggregation and irregular particle formation. These assembly issues will need to be remedied if the production of this VLP-based IAV vaccine is required to operate at a larger scale.

### 1.3.3 VLPs as Gene Therapy Products

Although the VLPs focused on in the project are predominantly vaccine candidates, there are many other therapies for which VLPs may be utilised. VLPs are being explored as possible treatments for rheumatoid arthritis and Alzheimer’s disease (Chackerian et al., 2001, Chackerian, 2010). Due to their ability to hold antigens on their surface as well as protect genetic material inside of their capsid shell, VLPs have been used to transport peptides, DNA, siRNA, and other small molecules (Shirbaghaee and Bolhassani, 2016). They have been employed as drug delivery mechanisms for a variety of treatments including anticancer antibiotics, chemotherapeutic drugs, and oncogene inhibitors (Zdanowicz and Chroboczek, 2016). Because they have the ability to contain genetic material and transduce it into cell genomes, VLPs have been explored as a possible route for gene therapy (Seow and Wood, 2009). Both JC polyomavirus VLPs and MPyV VLPs have been used to transport plasmid DNA and transfer it into mouse cells (Chao et al., 2016, Tegerstedt et al., 2005). Because adeno-associated virus (AAV) is one of the most common vectors for gene therapy, AAV VLPs are of particular interest as potential gene therapy products.

### 1.3.3.1 *Traditional Adeno-Associated Virus Gene Therapy*

Adeno-associated virus (AAV) is a small nonenveloped virus composed of three capsid proteins in a symmetrical icosahedral arrangement that contains a single-stranded DNA (ssDNA) genome of 4.7 kb. AAV is a promising candidate for gene therapies because it has the ability to transport short ssDNA sequences without integrating them into the host genome. There are currently over 50 clinical candidates for recombinant AAV therapies to treat an assortment of disorders with genetic markers including haemophilia, Alzheimer's disease, and Parkinson's disease (Naso et al., 2017). Two of the gene therapy products currently available, Spark Therapeutics' Luxturna and AveXis's Zolgensma, use AAV as the delivery vector.

AAV serotype 2 (AAV2) is the most well-studied serotype of the twelve AAV serotypes. The AAV2 genome consists of two open-reading-frame (ORF) cassettes. One ORF cassette codes for the non-structural proteins which aid in DNA replication, regulate transcription, target specific sites for gene integration and packaging of DNA into the capsid; these proteins are referred to as Rep 78, Rep 68, Rep 52, and Rep 40 (Ward et al., 2003, Pereira et al., 1997, Surosky et al., 1997, King et al., 2001). The other ORF cassette codes for the structural capsid proteins, VP1 (87 kDa), VP2 (73 kDa), and VP3 (62 kDa), and the assembly-activating protein (AAP) which promotes capsid protein interaction and improves stability (Rose et al., 1971, Maurer et al., 2018). The three capsid proteins are fairly similar in sequence; they are produced by alternative splicing so they only differ in their N-terminus (Becerra et al., 1988). The splicing and the leaky expression of the alternative reading frame for VP2 results in a molar ratio of 1:1:10 VP1:VP2:VP3 (Becerra et al., 1988).

The current method of recombinant AAV production uses triple transfection of a mammalian host cell, typically HEK293 (Figure 1.9). The host cell is first infected with a plasmid containing essential adenovirus genes to mediate AAV replication. Previously, an adenovirus would be introduced to the host cell to enable replication, but this requires that the adenovirus be purified out further downstream. By using a plasmid with the adenovirus genes, the presence of an additional virus can be circumnavigated altogether. Next, the host cells are transfected with a plasmid containing recombinant AAV genes to create the AAV itself and a plasmid for the gene therapy product. This method has been used to produce over 100 viral genomes per cell. However, production is limited by the fact that HEK293 cells are usually cultivated in adherent cultures. Currently, research is being undertaken to cultivate HEK293 cells in suspension in order to increase the volume of cells transfected and, consequently, the amount of recombinant AAV viral genomes produced (Blessing et al., 2018). This process also results in significant heterogeneity from process and product related impurities (Wright, 2014).

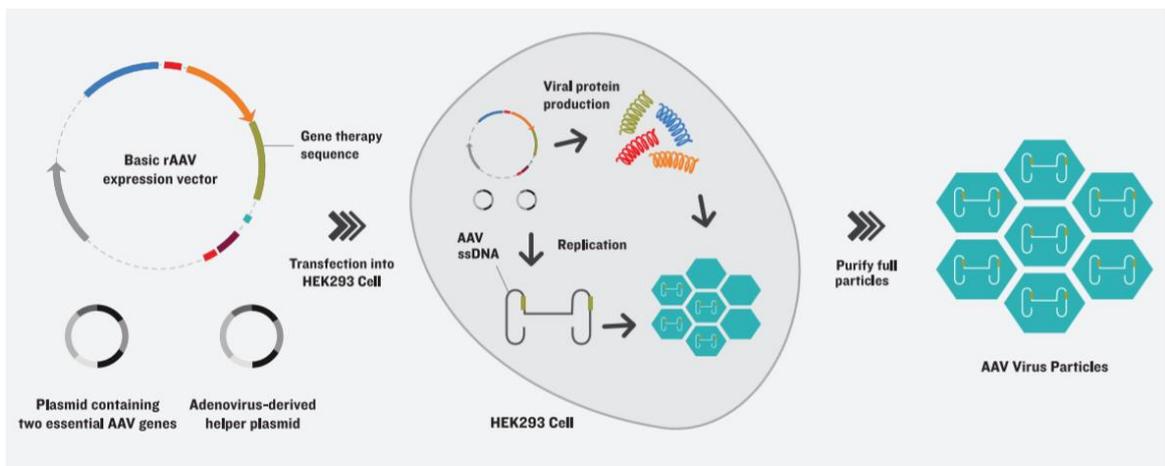


Figure 1.9 Overview of Triple-Transfection Strategy for Adeno-Associated Virus Production from Vigene Biosciences 2019

“This approach eliminates the need for a ‘helper’ adenovirus by introducing key genes in a DNA plasmid (top), accompanied by two other plasmids. The three plasmids are transfected into the production cell line (middle), which produces mature AAV capsids containing the transgene of interest. Once purified (bottom), these viral particles can infect patient cells to deliver their DNA payload but cannot replicate further.” Used with permission from (Biosciences, 2019)

The recombinant AAV gene therapy product is introduced to the patient through local intramuscular injection, systemic delivery, central nervous system delivery, cardiac delivery, or pulmonary delivery (Naso et al., 2017). The exact mechanism of gene transfer is summarised in the literature [Reviewed in (Wang et al., 2019a)]. Briefly, the recombinant virus binds to a glycosylated cell receptor and is internalised into an endosome which traffics the virus to the nucleus. The cargo is released, converted to double-stranded DNA (if it is not double-stranded already), and undergoes transcription. This genetic material will remain in the nucleus as circularised episomal genomes or it may integrate into the host cell genome.

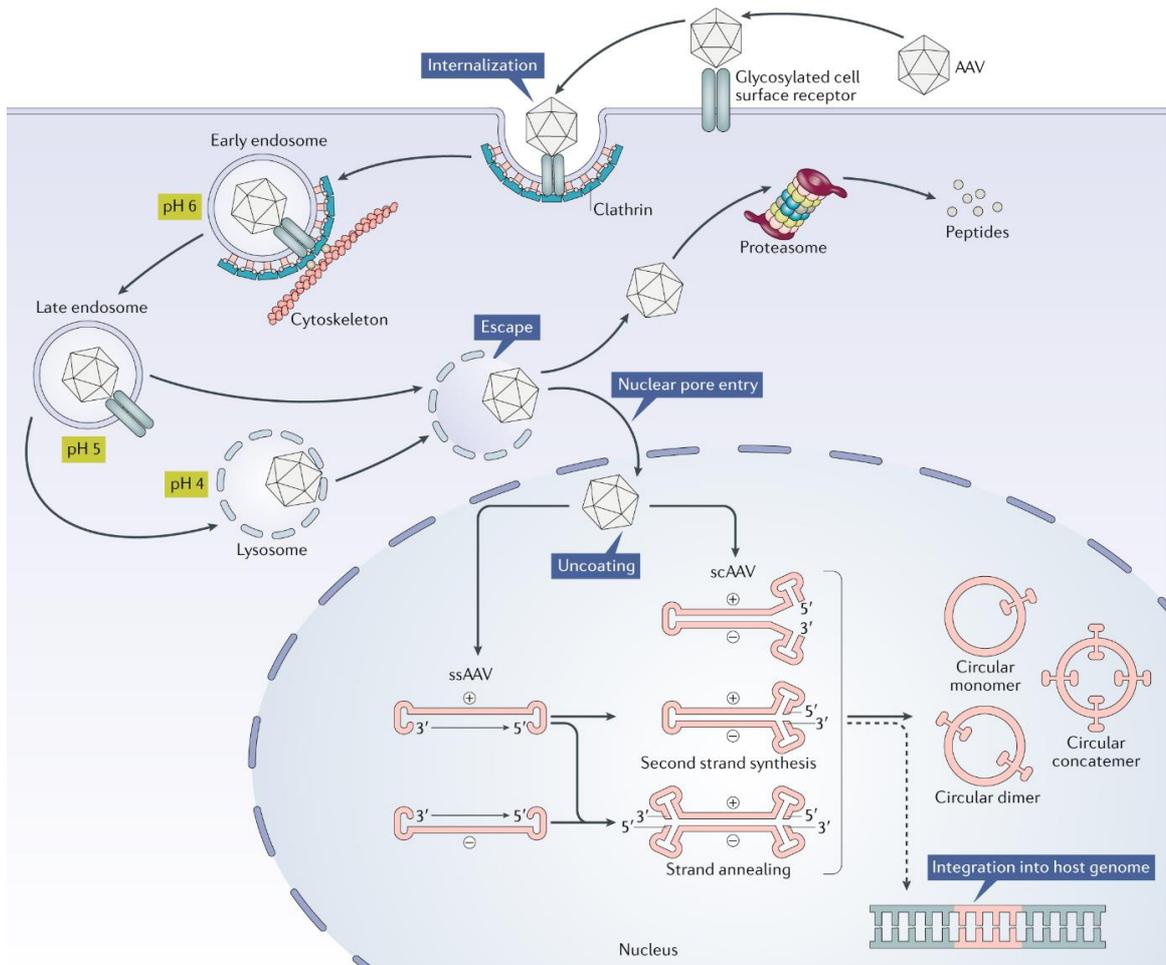


Figure 1.10 Diagram of Recombinant Adeno-Associated Virus Transduction Pathway from Wang et al, 2019a

In order to transfer the genetic material that it carries to a patient, AAV binds to glycosylated cell surface receptors. It enters the cell via endocytosis and is trafficked through the endosome to the cell's nucleus. There the virus undergoes uncoating and its cargo is unloaded. The virus may be a single-stranded AAV (ssAAV), in which case its genetic material must be converted to double-stranded DNA prior to transcription, or a self-complementary AAV (scAAV), in which case its genetic material is already double-stranded. This genetic material will form circularized episomal genomes that remain in the nucleus or it will be integrated into the host genome. If the AAV does not reach the nucleus, it will undergo proteolysis by the proteasome. Used with permission from (Wang et al., 2019a).

### 1.3.3.2 Adeno-Associated Virus Capsid Protein VLPs

AAV VLPs are empty shells formed only from the three capsid proteins. The first AAV VLPs assembled *in vitro* were formed from capsid proteins expressed by baculoviruses and purified separately that were then combined in a HeLa extract (Steinbach et al., 1997). AAV VLPs have also been produced in HEK293, *Saccharomyces cerevisiae*, and *E. coli* cells (Nieto et al., 2012, Backovic et al., 2012, Le et al., 2019). Mempro™, a Creative Biostructure company, currently

manufactures AAV VLPs using a mammalian expression system (Biostructure, 2020).

AAV VLPs have been used for a variety of applications as both a scaffold for antigen display for vaccination and as a delivery mechanism for genetic material (Nieto et al., 2012). AAV VLPs have been filled with ssDNA with the help of the Rep52 and Rep40 proteins (King et al., 2001). They have also been demonstrated to transport siRNA to cancer cells (Shao et al., 2012). Additionally, empty capsids can be formed from VP3 alone, although VP1 is necessary for the formation of infectious capsids (Smuda and Carter, 1991). AAV VP3 VLPs have been modified both on the C-terminus with foreign peptides and on the N-terminus with a FLAG tag for easier purification (Le et al., 2019, Hoque et al., 1999). Antigen display on AAV VLPs could open the door not only to AAV VLP-based gene therapy products, but more targeted therapies in general.

While AAV VLPs have their advantages for packaging and delivery of genetic material, traditional AAV viruses are currently more desirable because they are easier to produce and better able to reach the cell nucleus (Le et al., 2019). Additionally, *in vitro* assembly is a difficult process that requires not only the correct ratio of the three capsid proteins but optimal temperature, pH, ionic strength, and presence of additives for proper assembly (Le et al., 2019). In an open CFPS system, these properties may be more easily monitored and manipulated as needed.

#### 1.3.4 CFPS of VLPs

Cell-free protein synthesis has the potential to greatly improve VLP production. Because CFPS reactions are open, the reaction mixture can be modified to allow for conditions not typically used *in vivo*, like extreme pH levels or higher levels of ionic strength which may result in improved VLP assembly, or

improved redox conditions that may allow for better disulphide bond formation (Ceres and Zlotnick, 2002, Bundy and Swartz, 2011). CFPS would also allow for non-standard amino acid click-chemistry to attach antigens to the surface of the VLPs (Patel and Swartz, 2011).

While CFPS platforms are growing as a method of recombinant protein production, there has been little work done in the realm of VLP expression via CFPS. Over a decade ago, the first *E. coli*-based CFPS platform was used to produce VLPs; MS2 bacteriophage coat protein VLPs and truncated HBcAg VLPs were expressed in 30  $\mu$ L and 1 mL culture volumes with titres of approximately 400  $\mu$ g/mL in 12 hours for both species (Bundy et al., 2007). More recently, human norovirus-like particles were produced in an *E. coli*-based cell-free protein synthesis platform at roughly 600  $\mu$ g/mL in 4 hours and the encephalomyocarditis virus was synthesised in a mammalian cell-based CFPS system at a rate of 1000 plaque forming units (pfu)/mL over the course of 24 hours (Sheng et al., 2017, Kobayashi et al., 2012). These experiments have only occurred on a very small scale – in volumes of less than 1 mL. While these small volumes allow for rapid screening of products, in order to use CFPS as a production method, larger volumes will need to be explored and optimised.

### **1.4 Aims and Objectives**

The main aim of this project was to establish a novel manufacturing system that could be used to produce self-assembling particles for vaccine production and gene therapy in under a day. Vaccines are critical in the management of existing and emerging diseases. In pandemic situations, they will need to be developed and produced rapidly. The advent of gene therapy products (and other personalised and stratified medicines) will require manufacturing systems that not only result in high titres in short timescales but are flexible and robust. This will allow for the consistent production of a variety of products.

In order to achieve this goal, the following objectives were accomplished:

1. **Develop an on-demand manufacturing system at UCL.** This system needed to be rapid, robust, and flexible; therefore an *Escherichia coli*-based cell-free protein synthesis (CFPS) platform was chosen. This was the first CFPS platform to be developed at UCL. A variety of cell extracts from different *E. coli* strains, concentrated reaction mixtures, and plasmid designs were created and tested to determine the conditions that allowed for the highest titres. The most productive reaction conditions were used to demonstrate consistent scale up.
2. **Design a process development strategy for the manufacturing system.** Once the *E. coli*-based CFPS platform was developed, a deeper understanding of the critical process parameters in the system was needed to enable the design of a process development strategy. There is currently a lack of process development protocols for CFPS systems. . Critical process parameters (the concentration of the plasmid, the amount of the cell extract, the pH of the concentrated reaction mix, the temperature of the reaction, and the length of the reaction) were manipulated and analytical

tools like multivariate data analysis (MVDA) and design of experiments (DoE) were used to rapidly improve titre. Using the information gathered and the tools mentioned above, a process development strategy specific to CFPS was designed.

- 3. Demonstrate production of vaccines and gene therapy products in the *E. coli*-based CFPS system.** Model self-assembling products were used to develop a robust manufacturing system that could be used with a variety of newly developed biologic medicines. An ideal model product would be a complex, multi-domain product capable of displaying antigens or delivering genetic material. Virus-like particles (VLPs) were chosen because they fit the criteria and they have applications as both vaccine and gene therapy products. The tandem-core hepatitis B core antigen (HBcAg) VLPs developed as an attempt at a universal influenza vaccine and on adeno-associated virus serotype 2 (AAV2) VLPs which have been shown to transport genetic material potentially as a new form of gene therapy were the main focus of this project (Shao et al., 2012, King et al., 2001). In addition to using the CFPS platform to produce tandem-core HBcAg VLPs and AAV2 VLPs, product quality, specifically product related impurities, were observed and new iterations of the products were rapidly designed and produced.

## **2. Materials and Methods**

### **2.1 Chemicals and Media**

All chemicals were purchased from Sigma Aldrich (Dorset, UK) unless otherwise noted. Milli-Q water was used for buffer and media preparation.

### **2.2 *E. coli* cell extract strains and cell bank preparations**

The following *E. coli* strains were used for extract preparation: BL21 (DE3) (Thermo Fisher Scientific, Paisley, UK), BL21 Star™ (DE3) (Thermo Fisher Scientific), Rosetta™ (DE3), ClearColi® BL21 (DE3) (Lucigen, Middleton, WI, USA), and SHuffle T7 (New England Biolabs, Ipswich, MA, USA). Cell banks were prepared for each strain. LB-agar plates were prepared by pouring ~20 mL of freshly autoclaved LB-agar (10 g/L peptone, 10 g/L yeast extract, 10 g/L sodium chloride, and 15 g/L agar) into a 90 mm Sterilin™ plastic Petri dish from Thermo Fisher Scientific and allowing the LB-agar to cool for at least 1.5 hours at room temperature. The *E. coli* aliquots purchased were spread on the LB-agar plates and incubated overnight at 37°C in a Galaxy R CO<sub>2</sub> incubator from Eppendorf (Stevenage, UK). A colony was selected from the plate and incubated in 5 mL of Lysogeny broth (LB) media (10 g/L peptone, 10 g/L yeast extract, and 10 g/L sodium chloride) in a 50 mL Greiner conical tube (Greiner Bio-One, Kremsmünster, Austria) for ~16 hours in a Kühner ISF1-X Climo-Shaker shaking incubator (Basel, Switzerland). The cultured cells were then combined in a 1:1 ratio with sterilised 80% v/v glycerol in 1.5 mL microcentrifuge tubes from Star Labs (Milton Keynes, UK) and frozen at -80°C for future use.

### **2.3 Plasmid Design and Preparation**

#### **2.3.1 GFP Plasmids**

Commercial CFPS kit suppliers recommend using a plasmid which has been optimised for cell-free expression although plasmids typically employed for

cell-based expression can be used as well (2011). Two green fluorescent protein (GFP) plasmids were selected: pJL1, a superfolder GFP (sfGFP) plasmid, pJL1 was a gift from Michael Jewett (Addgene plasmid # 69496 ; <http://n2t.net/addgene:69496> ; RRID:Addgene\_69496), and pET14b-GFP, a plasmid developed by Martin Warren's group at the University of Kent used for *E. coli* expression of GFP+ with a 6x histidine tag. Plasmid pJL1 has been optimised for CFPS. It is a much smaller plasmid (2486 bp) that contains only the gene of interest (sfGFP in this case), the T7 promoter, the T7 terminator, a gene for kanamycin resistance and an origin of replication. The pET14b-GFP plasmid has not been optimised for cell-free protein synthesis and has not been codon-optimised for *E. coli*. It is also somewhat large compared to the pJL1 plasmid, 5389 bp.

### 2.3.2 HBcAg VLP Plasmids

All VLP plasmids were received from iQur Ltd. (London, UK). They were all designed for *in vivo* production. The monomeric construct was obtained from a lab at the Latvian Biomedical Research and Study Centre that is partnered with iQur Ltd. It is a plasmid for HBcAg subtype *ayw* under the T7 promoter in a pETDuet-1 backbone. Its exact sequence is unknown, but it is ~5900 bp long. The protein sequence for the monomeric construct is shown in Appendix A.

All of the tandem-core constructs are serotype *adw*. They are in kanamycin resistant pET28b plasmids. The K1K1 VLP is a hepatitis B tandem-core protein VLP with a single lysine in each MIR. It was received as a pPICZA plasmid, which is typically used for expression in *P. pastoris*. In order to keep the new K1K1 VLP plasmid as similar as possible to the other tandem-core plasmids, the K1K1 gene was codon-optimised for *E. coli* production and inserted into a pET28b vector; this plasmid was 6412 bp long. The protein sequence for K1K1 is found in Appendix A.

Once the K1K1 VLP plasmid synthesis was completed, it was transformed into chemically competent JM109 cells from Promega (Madison, WI, USA) and cell banks were prepared as explained above.

Plasmids for VLP 3, which has three matrix 2 ectodomain protein (M2e) variants in one MIR while the other MIR is empty, and VLP 1, which has a haemagglutinin stalk protein (HA2) in one MIR and three variants of M2e in the other MIR, were also prepared. These plasmids were 6637 bp and 6847 bp long respectively. Annotated versions of these sequences and the modified sequences mentioned in Chapter 5 can be found in Appendix A.

### 2.3.3 AAV2 Capsid Protein Plasmids

Plasmids were prepared for the three AAV2 capsid proteins called VP1, VP2, and VP3. The gene sequence for each plasmid was optimised for *E. coli* expression and inserted into the pJL1 vector. They were prepared both with and without N-terminal 6x histidine tags. The plasmid expressing VP1 was 4007 bp long (3971 bp without the 6x histidine tag). The plasmid expressing VP2 was 3596 bp long (3560 bp without the 6x histidine tag). The plasmid expressing VP3 was 3401 bp long (3365 bp without the 6x histidine tag). These sequences can be found in Appendix A.

### 2.3.4 Preparation of Plasmids for Cell-Free Protein Synthesis Reactions

The Qiaprep Spin Miniprep Kit and the HiSpeed Plasmid Maxi Kit, both from Qiagen™ (Manchester, UK), were used to prepare plasmids for CFPS reactions. The protocol for the Qiaprep Spin Miniprep Kit was followed as specified in the kit with the exception of the final step: 35 µL of elution buffer was used instead of the recommended 50 µL in order to better concentrate the DNA. When preparing cells to be used for the Qiaprep Spin Miniprep Kit, 50 µL of previously prepared cell bank containing the plasmid of interest and the appropriate antibiotic depending on the

plasmid (either 100 µg/mL ampicillin or 50 µg/mL kanamycin) vector were added to 5 mL of LB media in a 50 mL Greiner conical tube and incubated at 37°C and 250 rpm overnight. For the HiSpeed Plasmid Maxi Kit, two overnight cultures were prepared each with 50 µL of previously prepared cell bank containing the appropriate plasmid, antibiotic depending on the plasmid vector, and 2 mL LB media in a Greiner conical tube. In the morning, the 2 mL cultures were transferred to two 250 mL baffled glass shake flasks containing antibiotic and 300 mL LB media. The 300 mL cultures were incubated for 24 hours before being harvested for the HiSpeed Plasmid Maxi Kit. Both the 2 mL overnight cultures and the 300 mL cultures were incubated at 37°C and 250 rpm. The HiSpeed Plasmid Maxi Kit was performed as specified in the kit protocol with the volumes of buffers P1, N2, and N3 doubled. For both preparation protocols, the final DNA concentration was determined by placing 1 µL on the platform of a Nanodrop™ 2000c Spectrophotometer from Thermo Fisher Scientific and reading the concentration based on the absorbance at 260 nm.

## **2.4 CFPS Reactions**

### **2.4.1 Cell-Free Extract Preparation**

The extracts were derived from the BL21 (DE3) (Thermo Fisher Scientific), BL21 Star™ (DE3) (Thermo Fisher Scientific), Rosetta™ (DE3), ClearColi® BL21 (DE3) (Lucigen,), and SHuffle T7 (New England Biolabs) *E. coli* strains using the method outlined previously (Figure 2.1) (Hong et al., 2015). Briefly, a small volume, approximately 100 µL, of bacterial glycerol stock was used to inoculate 50 mL fresh LB medium pH 7.4 in a 250 mL baffled shake flask. The cultures were incubated overnight at 34°C (37°C for the ClearColi® BL21 (DE3) strain) and 250 rpm. The following morning, approximately 16 hours later, 25 mL of the overnight culture was transferred to 500 mL of 2xYTPG medium pH 7.2 (16 g/L tryptone, 10 g/L yeast

extract, 5 g/L NaCl, 7 g/L K<sub>2</sub>HPO<sub>4</sub>, 4.3 g/L KH<sub>2</sub>PO<sub>4</sub> and 18 g/L glucose; adjusted pH to 7.2 with potassium hydroxide) in a 2 L baffled shake flask. The culture was incubated at 34°C (37°C for the ClearColi® BL21 (DE3) strain) and 220 rpm until OD<sub>600</sub> ≈ 2 was achieved at which point 500 µL of 1 M potassium hydroxide was added to prevent acidification of the culture (as recommended by Hong et al. 2015) and the incubation continued (Hong et al., 2015). When OD<sub>600</sub> ≈ 4 was achieved, the culture was harvested by centrifugation at 5,000g and 4°C for 15 minutes; the cells should be entering stationary phase. Contrary to popular methods of harvesting during the mid-late log phase, Failmezger et al. showed that high performing extract can be produced from *E. coli* in the stationary phase; as such, their method of extract production was adopted to simplify the workflow (Failmezger et al., 2017).

The culture was then harvested by centrifugation at 5,000g and 4°C for 15 minutes using JA-10 rotor in a Beckman Coulter Avanti® J-20XP centrifuge (High Wycombe, UK). The supernatant was discarded and the pellets were kept on ice whenever possible. Each pellet was washed with ~25 mL of S30 buffer (pH 8.2 10 mM Tris acetate, 14 mM magnesium acetate, 60 mM potassium acetate, and 1 mM dithiothreitol) and resuspended by vortex. The resuspended pellet was pelleted by centrifugation at 9,000g and 4°C for 10 minutes in an Eppendorf 5910R centrifuge. The pellet was washed, resuspended, and pelleted by centrifugation again. Excess supernatant was discarded. Pellets were stored at -80°C following this step. Pellets were resuspended in 1.0 mL of S30 buffer per 1.0 g of pellet. The pellet was thawed on ice with S30 buffer for at least one hour prior to resuspension. The resuspended cells were homogenised via single pass at 1000 bar through an APV Gaulin Micron Lab40 Homogeniser (Lubeck, Germany). The homogenised lysate was clarified by centrifugation at 30,000g and 4°C for 30 minutes using a

JA-17 rotor in a Beckman Coulter Avanti® J-20XP centrifuge. (The speed of this centrifugation step is what give S30 lysates their name.) The supernatant was recovered and pelleted by centrifugation again. The supernatant from the second centrifugation was decanted and separated into 1 mL and 200 µL aliquots in Star Labs microcentrifuge tubes. Aliquots were flash frozen with liquid nitrogen and stored at -80°C until use.

The extract preparation method was not optimised or varied for different strains, as no evidence of any strain or product specific impact from extract preparation methods was found in the literature. While this protocol was chosen based on the equipment and reagents available, it is important to recognise that numerous groups have thoroughly examined the extract preparation protocol and optimised conditions such as length of time for cell growth, length of time for induction, and media for cell culture (Dopp and Reuel, 2018, Wilding et al., 2019, Zawada and Swartz, 2006, Levine et al., 2019, Bremer and Dennis, 2008).

In addition to the extracts prepared above, the SHuffle T7 extract and another BL21 Star™ (DE3) extract were prepared following the protocol above with the addition of 500 µL of 1 M IPTG when the 500 mL shake flask cultivation had achieved  $OD_{600} \approx 0.6$ . This induction was done to increase the amount of T7 RNA polymerase in the extract which has been demonstrated to improve expression (Des Soye et al., 2019). Unlike the other strains, these cells were incubated at 37°C.

A Bradford assay was used to determine total protein concentration for each extract; all extracts gave values of 30-50 mg/mL as expected based on previous literature (Kwon and Jewett, 2015, Caschera and Noireaux, 2015a). The total protein concentrations were as follows: 38 mg/mL for BL21 Star™ (DE3), 35 mg/mL for the BL21 (DE3), 34 mg/mL for Rosetta™ (DE3), 41 mg/mL for

ClearColi® BL21 (DE3), 40 mg/mL for IPTG-induced SHuffle T7, and 44 mg/mL for the IPTG-induced BL21 Star™ (DE3).

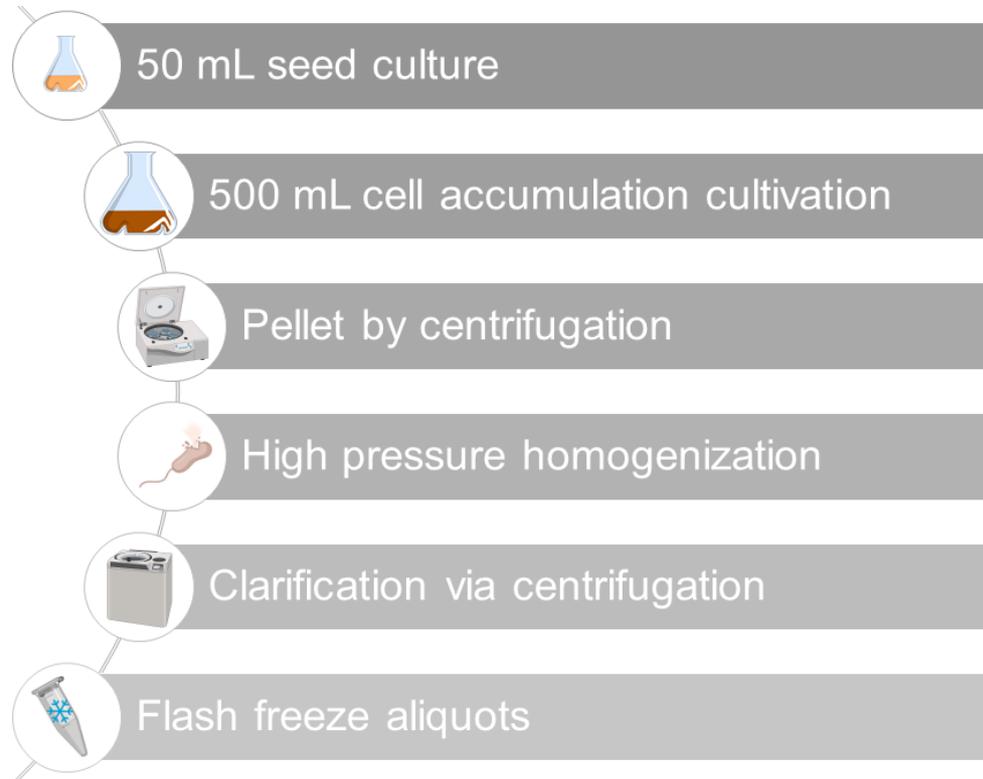


Figure 2.1 Cell Extract Preparation

Extract preparation involves cell growth, lysis via high pressure homogenisation, clarification, and storage. Created with Biorender.com.

#### 2.4.2 Cell-Free Concentrated Reaction Mixture Preparation

A cell-free concentrated reaction mixture based on the protocol used by Kwon and Jewett (2015) often referred to as the PANOx-SP system was prepared (Kwon and Jewett, 2015). The cell-free reaction included the following: 1.2 mM ATP, 0.85 mM each CTP, GTP, UTP, 1.5 mM spermidine, 1 mM putrescine, 33 mM phosphoenolpyruvate (PEP), 4 mM sodium oxalate, 0.27 mM coenzyme A (CoA), 0.33 mM nicotinamide adenine dinucleotide (NAD), 34 µg/mL folinic acid, 170 µg/mL tRNA from *E. coli* MRE 600, 90 mM potassium glutamate, 10 mM ammonium glutamate (MP Biomedicals, Eschwege, Germany), and 12 mM magnesium glutamate (Kwon and Jewett, 2015). A highly similar mixture has

previously been used with BL21 Star™ (DE3) based extracts and K12 MG1655 (strain C495) based extracts to achieve ~500 µg/mL sfGFP in 4 hours (Kwon and Jewett, 2015).

Another concentrated reaction mixture considered to be a simplified minimal mixture based on the protocol by Cai et al. (2015) was also prepared (Cai et al., 2015). This mixture included the following: 1.2 mM AMP, 0.86 mM each CMP, GMP, UMP, 1.5 mM spermidine, 4 mM potassium oxalate, 15 mM phosphate buffer pH 7.0, 260 mM potassium glutamate, 8 mM magnesium glutamate, and 2 mM oxidised glutathione (Cai et al., 2015). This mixture is based on what is often referred to as the “Cytomim” mixture because it was originally designed to mimic the cytosol environment in *E. coli* cells (Jewett and Swartz, 2004). Instead of using expensive phosphorylated energy sources, this mixture relies on oxidative phosphorylation and avoids the accumulation of inorganic phosphate and dramatic shifts in pH level over the course of the reaction (Jewett and Swartz, 2004, Jewett et al., 2008). The optimised version of this mixture employed by Cai et al. (2015) reduced reagent costs by 95% while still producing over 1 g/L trastuzumab single-chain fragment variable (scFv) (Cai et al., 2015).

Each concentrated reaction mixture was prepared as a 2.5x concentrated solution (without amino acids or T7 RNA polymerase). Potassium hydroxide was added to the complex concentrated reaction mixture to solubilise the components. 2.5x concentrated reaction mixture solutions were stored at -80°C until use. A 75 mM methionine solution and a 50 mM solution containing the remaining nineteen amino acids were prepared separately and added to the reaction to a final concentration of 1.25 mM of each amino acid except methionine for which the final concentration was 1.5 mM. 50 U/µL T7 RNA polymerase (Catalogue number: 18033019) was purchased from ThermoFisher Scientific™ and added to the

reaction to a final concentration 500 U/mL T7 RNA polymerase. Amino acids solutions and T7 RNA polymerase were stored at -20°C until use.

#### 2.4.3 Cell-Free Protein Synthesis Reaction Ratios

For a typical CFPS reaction, 10 µg/mL plasmid (depending on the plasmid used, the molar concentration will also be noted in the text), 20% v/v cell extract, 40% v/v concentrated reaction mixture, 2% v/v 75 mM methionine solution, 2.5% v/v 50 mM amino acid (-methionine) solution, and 1% v/v 50 U/µL T7 RNA polymerase (Thermo Fisher Scientific, Catalogue number: 18033019) were combined. Nuclease free water from Thermo Fisher Scientific was used to make up the remaining volume. For the first 30 minutes of the reaction, only the following were added: 10 µg/mL plasmid, 20% v/v cell extract, 20% v/v concentrated reaction mixture, 1% v/v 75 mM methionine solution, 1.25% v/v 50 mM amino acid (-methionine) solution, and 1% v/v T7 RNA polymerase solution. Reactions were incubated at 30°C for an additional 4.0 hours at 300 rpm in a shaking incubator or 1200 rpm in an Eppendorf Thermomixer® C. These conditions were used for all reactions unless otherwise specified. Reactions were stored for future use and analysis at -20 °C.

#### 2.4.4 Adjusting Process Parameters

Each of the following process parameters were examined in isolation in order to determine their impact on titre: plasmid concentration, amount of extract, temperature, pH of the concentrated reaction mixture, and length of reaction. Plasmid concentration was adjusted by altering the volume of concentrated plasmid added to each reaction. Final concentrations of 1.0 µg/mL, 5.0 µg/mL, 10 µg/mL, 20 µg/mL, and 50 µg/mL (0.61 nM, 3.1 nM, 6.1 nM, 12.2 nM, and 30.5 nM for sfGFP and 0.26 nM, 1.3 nM, 2.6 nM, 5.1 nM, and 12.8 nM for HBcAg) were prepared. Amount of extract was likewise adjusted by volume. Reactions with 5%

v/v, 10% v/v, 15% v/v, 20% v/v, 25% v/v, 30% v/v, and 35% v/v extract were prepared. The pH level of 2.5x complex concentrated reaction mixture was adjusted using small volumes of 18 M hydrochloric acid or 12 M sodium hydroxide. Below pH 5.5, the components began to fall out of solution. The following pH levels were tested: 5.0, 5.5, 6.0, 6.5, 6.8, 7.0, 7.5, 8.0, 8.5, and 9.0. Temperature was adjusted on each thermomixer; reactions at 15°C, 20°C, 25°C, 30°C, 32°C, 35°C, 37°C and 40°C were tested. The following reaction lengths were assessed: 0.5 hours, 1 hour, 2 hours, 4 hours, 6 hours, 8 hours, 12 hours, 22 hours, and 24 hours. For sfGFP, the typical reactions conditions were 6.1 nM (10 µg/mL) plasmid, 20% v/v non-induced BL21 Star™ (DE3) extract, complex concentrated reaction mixture at pH 6.8, 20°C, and 4 hours. Reactions adjusting the parameters were performed in triplicate and a total of 105 reactions were performed: typical reactions conditions (9 reactions), plasmid concentration at 4 levels (12 reactions), amount of extract at 6 levels (18 reactions), pH at 8 levels (24 reactions), temperature at 7 levels (21 reactions), and reaction length at 7 levels (21 reactions). Based on the sfGFP results, some reaction conditions were omitted for HBcAg. For HBcAg, the typical reactions conditions were 2.6 nM (10 µg/mL) plasmid, 20% v/v non-induced BL21 Star™ (DE3) extract, complex concentrated reaction mixture at pH 6.0, 20°C, and 4 hours. Reactions were performed in duplicate and a total of 48 experiments were performed: typical reaction conditions (2 reactions), plasmid concentration at 4 levels (8 reactions), amount of extract at 5 levels (10 reactions), pH at 4 levels (8 reactions), temperature at 4 levels (8 reactions), and reaction length at 6 levels (12 reactions).

#### 2.4.5 Expressway™ Cell Free Expression System

The CFPS platform as developed was compared to the Expressway™ Cell Free Expression System sold by ThermoFisher Scientific™. For a 100 µL reaction,

20  $\mu\text{L}$  *E. coli* slyD- extract, 20  $\mu\text{L}$  2.5x IVPS *E. coli* reaction buffer (concentrated reaction mixture), 1  $\mu\text{L}$  50 U/  $\mu\text{L}$  T7 RNA polymerase, 1  $\mu\text{L}$  75 mM methionine, 1  $\mu\text{L}$  50 mM amino acids (-methionine), and 1  $\mu\text{g}$  DNA template were combined in a microcentrifuge tube. Nuclease-free water was added to a final volume of 50  $\mu\text{L}$ . The tubes were incubated at 30°C and 300 rpm for 0.5 hours – for the Eppendorf Thermomixer® C, 300 rpm is achieved when the shake rate is set to 1200 rpm. After 0.5 hours incubation, 25  $\mu\text{L}$  2.0x IVPS *E. coli* reaction buffer (concentrated reaction mixture), 1  $\mu\text{L}$  75 mM methionine, 1  $\mu\text{L}$  50 mM amino acids (-methionine), and 22.75  $\mu\text{L}$  nuclease-free water were added to each tube for a final volume of 100  $\mu\text{L}$ . The tubes were incubated at 30°C and 300 rpm for an additional 4.0 hours (for a total of 4.5 hours). Reactions were analysed immediately and stored at -20°C.

#### 2.4.6 Vessels and Scales for CFPS Reactions

A range of reaction platforms were used in this project. 100  $\mu\text{L}$  reactions were performed in Star Labs microcentrifuge tubes agitated in an Eppendorf Thermomixer® C and Corning Costar® black well, clear flat bottom microwell plates (Amsterdam, Netherlands) agitated in a BMG Labtech (Aylesbury, UK) FLUOStar OPTIMA spectrophotometer. 1 mL reactions were performed in deep-well flower plates sold by m2p-labs (Baesweiler, Germany) agitated in an Eppendorf Thermomixer® C. 10 mL reactions were performed in 250 mL baffled shake flasks and Sarstedt tissue culture T-75 flasks (Nümbrecht, Germany) agitated at 200 rpm in a Kühner ISF1-X Climo-Shaker shaking incubator. 100 mL reactions were performed in an Eppendorf DASbox Mini Bioreactor System. An agitation rate of 500 rpm and a sparge rate of 1 vvm (6 L/h) was used in order to ensure that the reaction was well-mixed.

## **2.5 Analytical Methods**

### 2.5.1 Fluorescence Analysis to Determine GFP titre

For GFP analysis, it is assumed that all GFP proteins that have been produced are correctly folded and emit with the characteristic fluorescence intensity for that GFP variant. Titre was measured through fluorescence intensity measurement on a BMG Labtech FLUOStar OPTIMA spectrophotometer at an excitation wavelength of 485 nm and an emission wavelength of 520 nm and compared to a standard curve of rTurbo GFP from Evrogen (Moscow, Russia). The range of the standard curve was 80 µg/mL to 1.6 µg/mL. To dilute into this range, all CFPS reactions were diluted ten-fold [20% v/v non-induced BL21 Star™ cell extract diluted in Milli-Q water]. This dilution should also mitigate the difference in the pH of the CFPS reactions. For different GFP variants, fluorescence intensity was scaled based on quantum yield and extinction coefficient using the following equation where “F” is the measured fluorescence of a sample, “φ” is the quantum yield for that variant, “I<sub>0</sub>” is the intensity of the incident light, “ε” is the extinction coefficient for that variant, “l” is the optical path length, and “c” is the concentration of a sample.

$$F = \phi I_0 (1 - 10^{-\epsilon lc})$$

The quantum yields for rTurbo GFP, sfGFP, and GFP+ are 0.53, 0.65, and 0.72 respectively (Evrogen, 2020, FPbase, 2020b, FPbase, 2020a). The extinction coefficients for rTurbo GFP, sfGFP, and GFP+ are 70000 M<sup>-1</sup>cm<sup>-1</sup>, 83300 M<sup>-1</sup>cm<sup>-1</sup>, and 82,400 M<sup>-1</sup>cm<sup>-1</sup> respectively (Evrogen, 2020, FPbase, 2020b, FPbase, 2020a).

CFPS GFP samples were also analysed via sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Samples were reduced with NuPAGE Sample Reducing Agent (Thermo Fisher Scientific) and NuPAGE MES Sample Buffer (Thermo Fisher Scientific). The samples were not boiled because

that would completely denature the protein and destroy fluorescence; previously, groups have used this method to visualise fluorescent proteins in SDS-PAGE gels (Bird et al., 2015). Samples were diluted four-fold with Milli-Q water and the appropriate buffers and then applied to the lanes of a NuPAGE 12% Bis-Tris gel (Thermo Fisher Scientific) at 200V for 50 minutes. Gels were imaged under blue fluorescent light (460 nm) on the GE Amersham™ Imager 600 (Pittsburgh, PA, USA). The gels were then stained with InstantBlue™ Coomassie Protein Stain and imaged again under white light.

### 2.5.2 Total Protein Analysis with Bradford Assay

The protein concentration of a sample can be determined with a Bradford assay. The Coomassie Brilliant Blue G-250 dye in the Bradford reagent binds to the proteins in the sample which results in a blue colour change. When compared to standard curve of proteins of known concentration, the protein concentration of a sample can be calculated.

The Quick Start™ Bradford Protein Assay protocol and Quick Start™ Bradford 1x Dye Reagent from Bio-Rad (Hercules, CA, USA) were used. Bovine serum albumin standards were prepared at the following concentrations: 2.0 mg/mL, 1.0 mg/mL, 0.75 mg/mL, 0.50 mg/mL, 0.25 mg/mL, 0.125 mg/mL, and 0 mg/mL. 5 µL of each standard and sample were added to a single well on a clear 96-well plate. 250 µL of Quick Start™ Bradford 1x Dye Reagent was added to each well. The plate was incubated for 5 minutes at room temperature. The absorbance of each well at 595 nm was measured using a CLARIOStar® Plus plate reader from BMG LabTech. Each standard and sample was measured in triplicate.

### 2.5.3 SDS-PAGE Analysis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is a common method of protein separation that involves denaturing the proteins

with SDS, which also gives the proteins a negative charge, before placing them on a polyacrylamide gel and applying an electrical current. This causes the proteins to migrate through the gel toward the anode. Smaller molecules migrate faster than larger molecules allowing for separation based on molecular weight.

Samples were reduced with 4x SDS-PAGE sample loading buffer before being boiled at 90°C for 10 minutes. 4x SDS-PAGE sample loading buffer contains 50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1%  $\beta$ -mercaptoethanol, 12.5 mM EDTA, and 0.02% bromophenol blue. Samples were diluted two-fold with Milli-Q water and the appropriate buffer and then applied to the lanes of a NuPAGE 12% Bis-Tris gel at 200V for 50 minutes. The gels were then stained with InstantBlue™ Coomassie Protein Stain and imaged using the GE Amersham™ Imager 600.

#### 2.5.4 Dot Blot Analysis

Dot blots are useful in quickly determining the presence of the product in a sample, especially when several samples are being collected, like fractions from a chromatography run or an ultracentrifugation procedure. Dot blots are limited in that they only provided information about whether or not a species that corresponds to the primary antibody is present. Truncated products or product related impurities might also contribute to a positive signal on a dot blot.

A standard curve of recombinant HBcAg (ab49014) from Abcam (Cambridge, UK) was prepared at the following concentrations: 250  $\mu$ g/mL, 125  $\mu$ g/mL, 62.5  $\mu$ g/mL, 31.3  $\mu$ g/mL, 15.6  $\mu$ g/mL, and 0  $\mu$ g/mL. 2  $\mu$ L of each standard and sample were applied as a “dot” on an 8 mm x 8 mm nitrocellulose membrane. Standards and samples were measured in triplicate. The membrane was blocked in Tris-Buffered Saline with Tween 20 and milk powder (TBST-M) for 45 minutes. The membrane was incubated with the primary antibody (Table 2.1 shows the antibodies used with each product) diluted 1:1000 in TBST-M for two hours. The

membrane was washed with TBST three times before being incubated with the secondary antibody, an anti-mouse HRP-conjugated antibody [HAF007] (R&D Systems, Abingdon, UK), diluted 1:1000 in TBST-M for one hour. The membrane was washed with TBST twice and TBS once. The membrane was incubated with the Thermo Fisher Scientific Pierce ECL Western Blotting Substrate for 1-2 minutes in darkness then exposed and imaged using the GE Amersham™ Imager 600. The image was analysed using ImageQuant software and the quantification of the samples was determined with four-parameter logistic (4PL) fitting.

Table 2.1 Mouse Primary Antibodies for Dot Blots and Western Blots

<b>Antibody name</b>	<b>Corresponding region</b>
Anti-Hepatitis B Virus Core Antigen antibody [10E11] (ab8639)*	Amino acids 1-10 on HBcAg
Anti-Hepatitis B Virus Core Antigen antibody [14E11] (ab8638)*	Amino acids 135-141 on HBcAg
Anti-Influenza A Virus M2 Protein antibody [14C2] (ab5416)*	N-terminal of the Influenza A Virus M2 Protein
Anti-6X His tag® antibody [HIS.H8] (ab18184)*	Any 6x histidine tag
Anti-Adeno-associated Virus / AAV (VP1+ VP2 + VP3) Monoclonal Antibody (clone: B1)**	Amino acids 726-733 on all three AAV capsid proteins

\*from Abcam (Cambridge, UK) \*\*from Generon

### 2.5.5 Western Blot Analysis

In CFPS, where there are so many proteins present that SDS-PAGE gels provide little information on the product of interest, western blots are essential to determine the presence of a particular protein in the reaction. For a western blot, the proteins in the polyacrylamide gel are transferred to a membrane using an electrical current. The membrane is then blocked and incubated with a target specific antibody. Next, the membrane is incubated with a second antibody that is specific to the first antibody. The secondary antibody is conjugated to another

molecule that allows for easy detection, for example, biotin or horseradish peroxidase. This results in the detection of specific proteins at specific molecular weights. For a given product, there may be a variety of primary antibodies to choose from. By applying different antibodies to a single protein sample, the nature of cleavage products or product related impurities can be better understood.

While western blots provide an easy way to understand whether or not a particular protein is present, because the proteins are reduced in SDS-PAGE, this technique provides no information on the assembly of the particle. Furthermore, while densitometry can be performed with various known concentrations of the protein of interest in a western blot, this is an inexact method of determining product titre. Additional techniques are required for more definitive titre and assembly analysis.

SDS-PAGE was performed as stated previously. The gels were transferred to an 8 mm x 8 mm nitrocellulose membrane using the Bio-Rad Trans-Blot® Turbo™ Transfer System. After transfer, the membrane was stained with Ponceau S for 5 minutes, de-stained with Milli-Q water and imaged. The stain was removed with TBS. The membrane was blocked in TBST-M for 45 minutes. The membrane was incubated with the primary antibody diluted 1:1000 in TBST-M for two hours. The membrane was washed with TBST three times before being incubated with the secondary antibody, an anti-mouse HRP-conjugated antibody, diluted 1:1000 in TBST-M for one hour. The membrane was washed with TBST twice and TBS once. The membrane was incubated with the Thermo Fisher Scientific Pierce ECL Western Blotting Substrate for 1-2 minutes in darkness then exposed and imaged using the GE Amersham™ Imager 600.

### 2.5.6 Densitometry Analysis

A densitometry analysis of an SDS-PAGE analysis image can be used to estimate the size of each band compared to other bands in that same lane. The percentage of protein in each band can then be compared to the overall protein concentration determined via Bradford assay to approximate the concentration of protein in a single band on the SDS-PAGE analysis.

This technique was used to determine the titre of the tandem-core hepatitis B core antigen (HBcAg) virus-like particles (VLPs) in Chapter 6. An SDS-PAGE analysis and a western blot analysis were performed on each sample. The images from both analyses were then aligned to determine which bands on the SDS-PAGE corresponded to the full-length product as shown on the western blot. The densitometry of the SDS-PAGE image was analysed using ImageQuant software to determine the percentage of protein in each of the previously identified bands. That percentage was applied to the results of a Bradford assay performed on each sample to calculate titre of that sample.

### 2.5.7 ELISA

Enzyme-linked immunosorbent assays (ELISAs) can be used to determine the titre of a specific protein within a sample. ELISAs are performed in 96-well plates. A capture antibody specific to the protein of interest is bound to the bottom of the 96-well plate. The sample containing the protein of interest is applied and incubated. A secondary antibody specific to the protein of interest is added to the 96-well plate. This secondary antibody may be conjugated to an enzyme or a third enzyme-conjugated antibody that binds the secondary antibody may be added. A substrate is added that is converted by the enzyme to create an absorbance, fluorescence or luminescence signal that can be detected and correlated against a standard curve of the protein of interest of known concentrations.

Because ELISAs are very specific, this technique was used in this project to determine yield even in a crude CFPS reaction sample. However, there is no way to determine the proportion of monomers, dimers, aggregates, and assembled particles in the reaction. The reaction sample would need to be purified using another analytical technique before an ELISA could be used to determine the titre of assembled product.

For the HBcAg VLPs, the QuickTitre™ Hepatitis B Core Antigen (HBVcAg) ELISA Kit from Cell Biolabs (San Diego, CA, USA) was used according to the recommended protocol. The exact antibodies used in this protocol are unknown so it is highly likely that some form of product related impurities are contributing to the reported yields.

#### 2.5.8 Ammonium Sulphate Precipitation of VLPs

The solubility of particles in a solution can be decreased by increasing the ionic strength of that solution. This causes the particles to become insoluble and precipitate. Ammonium sulphate is typically used to precipitate particles because it is highly soluble and because the ions it disassociates into, ammonium and sulphate, have been shown to stabilise protein structure (2014).

The VLPs produced in this project were purified via ammonium sulphate precipitation. Either the soluble fraction or the resuspended pellet (also referred to as the insoluble fraction) obtained from the reaction was incubated with 1.9 M ammonium sulphate for 5 minutes. After the particles were pelleted by centrifugation at 15,000g for 10 minutes, they were resuspended in renaturing buffer, 0.1 M tris buffer pH 8.7, 1 mM EDTA, 0.15 M NaCl.

#### 2.5.9 Transmission Electron Microscopy

Transmission electron microscopy (TEM) is an analytical technique in which an image is formed by electrons being transmitted through an ultrathin sample.

TEM allows for the observation of particles at sub-nanometre resolutions. It is the only way to tell with absolutely certainty that the virus-like particles have assembled correctly. Several authors have demonstrated assembly of their VLP products using TEM imaging (Bundy et al., 2007, Holmes et al., 2015, Peyret et al., 2015). But TEM is not without its concerns. Samples need to be purified and dried before they can be examined on the microscope which may distort particles and make it more difficult to accurately determine the concentration and size of the particles. TEM is one the best techniques available to observe particle assembly, but it should be compared with other techniques to ensure that the precision of the information obtained.

The samples were applied to a carbon/formvar-coated copper 300 mesh grids purchased from Geron (Slough, UK) for 1 minute. The grid was washed with water for 5 seconds and then negatively stained with 2% v/v uranyl acetate in water for 30 seconds. The grids were imaged at UCL with the assistance of Mark Turmaine under a JEOL JEM-1010 transmission electron microscope (Welwyn Garden City, UK) and imaged under a Gatan Orius camera (Abingdon, UK). The grids were imaged at the University of Kent with the assistance of Ian Brown using a JEOL JEM-1230 microscope and imaged under a Gatan multiscan digital camera.

#### 2.5.10 Nickel-Nitrilotriacetic Acid Resin Binding for Histidine-Tagged Proteins

Nickel-nitrilotriacetic acid (Ni-NTA) resin is an agarose resin charged with Ni<sup>2+</sup> ions. It binds proteins with a 6x histidine tag. These proteins can then be eluted from the resin by decreasing the pH or adding a competing substance that also binds to the resin, like imidazole.

1 mL of Ni-NTA resin in ethanol (Thermo Fisher Scientific) was transferred to a 2 mL Star Labs microcentrifuge tube. The resin was pelleted by centrifugation

at 7000 rpm for 2 min in an Eppendorf 5427 R centrifuge. The liquid was decanted and 1 mL of Milli-Q water was added to the tube. The tube was inverted several times. The resin was pelleted by centrifugation at 7000 rpm for 5 min. The liquid was decanted and 0.5 mL binding buffer (20 mM Tris-HCl, 500 mM NaCl, 5 mM imidazole, pH 8.0) was added to the tube. The tube was inverted several times. The resin was pelleted by centrifugation at 7000 rpm for 2 min. The liquid was decanted and 0.5 mL binding buffer was added to the tube. The tube was inverted several times. The resin was pelleted by centrifugation at 7000 rpm for 5 min. The liquid was decanted and 0.4 mL of protein in Renaturing Buffer (0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7) was added to the tube. The resin and the sample were incubated at 4°C overnight on a shaking platform.

In the morning, approximately 16 hours later, the resin was pelleted by centrifugation at 7000 rpm for 2 min. The liquid was decanted and 0.5 mL binding buffer was added to the tube. The tube was inverted several times and incubated at 500 rpm, room temperature, for 5 min. The resin was pelleted by centrifugation at 7000 rpm for 2 min. The liquid was decanted and 0.8 mL wash buffer 1 (20 mM Tris-HCl, 500 mM NaCl, 50 mM imidazole, pH 8.0) was added to the tube. The tube was inverted several times and incubated at 500 rpm, room temperature, for 5 min. The resin was pelleted by centrifugation at 7000 rpm for 2 min. The liquid was decanted and 0.8 mL wash buffer 2 (20 mM Tris-HCl, 500 mM NaCl, 100 mM imidazole, pH 8.0) was added to the tube. The tube was inverted several times and incubated at 500 rpm, room temperature, for 5 min. The resin was pelleted by centrifugation at 7000 rpm for 2 min. The liquid was decanted and 0.4 mL elution buffer (100 mM Tris, 1 mM EDTA, 500 mM NaCl, 250 mM imidazole, pH 8.7) was added to the tube. The tube was inverted several times and incubated at 500 rpm, room temperature, for 5 min. The resin was pelleted by centrifugation at 7000 rpm

for 2 min. The liquid was decanted. All decanted liquid was stored in separate microcentrifuge tubes. The samples and resin were stored at -20°C

### 2.5.11 Multivariate Data Analysis

MVDA was used to evaluate the results from the reactions detailed in section 2.4.4 to determine which combination of conditions would maximise titre and to gain a better understanding of each variable's contribution on titre for both products investigated. Data manipulation and analysis for the multilinear regression (MLR) model was performed using MATLAB R2019b (MathWorks, Inc., Natick, MA). MLR was used to predict a single dependent variable – titre – from a series of independent inputs – plasmid concentration, amount of *E. coli* extract in the reaction, pH of concentrated reaction mix, temperature of reaction, and length of reaction. In this manner, the contribution of each independent parameter towards titre can also be determined. The variables of importance were found by creating multiple MLR models that studied the influence of each parameter that was removed during the development of models that considered linear, quadratic, polynomial (squared terms and cubed terms), and interactions. Separate models were created for sfGFP and HBcAg titres. The prediction performance of the MLR was quantified using the coefficient of determination which is calculated as:

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

Where,  $y_i$  is the product concentration for run  $i$ ,  $\bar{y}$ , is the product concentration mean and  $\hat{y}_i$  is the predicted product concentration for run  $i$ . The MLR model terms were chosen based on a stepwise regression approach implementing both forward addition and backward elimination of terms based on their p-value which ensured a robust and statistically valid model. The selection criteria for the finalised model was based on maximising the coefficient of

determination between the model predictions and the experimental product concentrations.

#### 2.5.12 Design of Experiments for Process Parameter Analysis

Following the “one-variable-at-a-time” analysis detailed in the previous section, a DoE study was performed for each product to better understand the interactions between different process parameters and the subsequent impact on titre. This exercise was done to demonstrate how titre might be maximised using DoE and to validate its use as part of this process development strategy. It is not the intention to optimise titre, although that could be done by expanding the design space in subsequent DoE studies. A different DoE approach was used for each of the two products. For sfGFP, a face-centred central composite design consisting of 9 runs with 2 centre points was used. The following conditions were chosen: temperature – 32°C, 34°C, 36°C; pH of concentrated reaction mixture – 5.5, 6.0, 6.5. BL21 Star<sup>TM</sup> extract was used with 6.1 nM (10 µg/mL) plasmid and 20% v/v cell extract for 4.0 hours. For HBcAg, a face-centred central composite design consisting of 27 runs with 2 centre points was used. The following conditions were chosen: temperature – 32°C, 34°C, 36°C; plasmid concentration – 3.8 nM, 7.7 nM, 11.5 nM (15 µg/mL, 30 µg/mL, 45 µg/mL); amount extract – 15% v/v, 20% v/v, 25% v/v. BL21 Star<sup>TM</sup> extract was used with 10 µg/mL plasmid and 20% v/v cell extract for 4.0 hours. These reactions were performed with an induced BL21 Star<sup>TM</sup> extract with pH 7.0 concentrated reaction mixture for 4.0 hours.

### **3. Developing a CFPS Manufacturing System for On-Demand Production**

*The on-demand manufacturing system in this project was an E. coli-based cell-free protein synthesis (CFPS) system. In developing this CFPS system, the three major components - the cell extract, the concentrated reaction mixture, and the plasmid for the product of interest - were carefully considered. Five different E. coli strains for cell extract production were examined and it was determined that three of these strains resulted in similarly high yields. A complex concentrated reaction mixture using phosphorylated energy sources for ATP regeneration, rather than oxidative phosphorylation, gave higher yields. A shorter plasmid backbone optimised for CFPS production performed better than a backbone designed for in vivo production, but optimisation was not a requirement for achieving detectable titres. In the end, the CFPS system in this project consists of a BL21-Star<sup>TM</sup> (DE3) extract, a concentrated reaction mixture based on the protocol in Kwon and Jewett, 2015, and a plasmid expressing superfolder green fluorescent protein (sfGFP) in a backbone optimised for CFPS production. Scale up of this platform from 100  $\mu$ L to 100 mL was demonstrated. sfGFP production was also used to benchmark the activity of different batches of cell extract to monitor batch-to-batch variation. Now that this CFPS manufacturing platform has been developed and an understanding of the three major reaction components has been achieved, other process parameters can be manipulated to improve titre and a process development strategy can be designed.*

#### **3.1 Introduction**

The first objective in developing an on-demand manufacturing scheme for this project was to design a platform that uses cell-free protein synthesis (CFPS) to rapidly produce self-assembling particles. There are three major components in a CFPS platform: the cell extract, the concentrated reaction mixture, and the DNA

plasmid. An *E. coli*-based extract was chosen because of the large quantity of literature regarding *E. coli*-based extracts for CFPS and the relatively quick and cost-effective cultivations required. There are several *E. coli* strains to choose from, each with their own properties that may be advantageous for a CFPS platform. A typical *E. coli* expression strain, BL21 (DE3), and its variants, BL21 Star™ (DE3), Rosetta™ (DE3), and ClearColi® BL21 (DE3) which have been used previously for CFPS extracts were selected (Krinsky et al., 2016, Kwon and Jewett, 2015, Kim et al., 2006b, Wilding et al., 2019). The SHuffle T7 strain was also selected as a candidate due to the interest in disulphide bond formation in *E. coli* extracts. For the concentrated reaction mixture, there are a multitude of formulations based on different energy sources to sustain the CFPS reaction that can be chosen. A minimal mixture proposed by Cai et al. (2015) called the Cytomim system was examined because it is one of the more cost-effective mixes in the literature and it contains fewer components compared to other mixes. The reaction mixture used in Kwon & Jewett (2015) known as the PANOX-SP system, which has been used successfully in several other publications, was also chosen (Lu et al., 2015, Salehi et al., 2016, Smith et al., 2012).

Plasmid design and its potential impact on titre was also explored. CFPS reactions do not require that a full plasmid be used to generate protein. In fact, protein has been produced in CFPS reactions from a variety of genetic material including linear DNA and mRNA (Wu et al., 2007, Hansen et al., 2016). Plasmid DNA was selected as the source of genetic material for protein production because it is easy to generate large amounts of plasmid DNA *in vivo* and purify it using pre-made kits. Using a minimal plasmid, like the pJL1 plasmid designed by the Jewett Lab at Northwestern University, may streamline resources for transcription that

might be otherwise occupied when a plasmid for typical *in vivo* protein production is used.

Once the CFPS platform was designed, its effectiveness was demonstrated at a variety of scales. CFPS reactions are typically done at the sub-millilitre scale; however, Sutro Biopharma recently demonstrated consistent scaling from 250  $\mu\text{L}$  reactions in microcentrifuge tubes to 100 L reactions in bioreactors while achieving titres above 700  $\mu\text{g/mL}$  (Zawada et al. 2011). Although CFPS reactions are typically more expensive than traditional *in vivo* cultivations, proven scale up of CFPS reactions has two major benefits. First, by establishing linear scaling across reaction volumes, more confidence can be placed in scale down CFPS reactions. Due to the high throughput nature of CFPS, thousands, if not millions of reaction conditions can be tested in a very short timescale for any given product. Second, large scale CFPS reactions open the door to mass production of previously difficult-to-express products like membrane proteins, toxic proteins, and proteins that have been incorporated with nonstandard amino acids (nsAAs) that might not otherwise have found a viable path to market.

### **3.2 Comparing Commercially Available Strains for Cell Extract**

Five different commercially available *E. coli* strains for the cell extract were examined: BL21 (DE3), BL21 Star<sup>TM</sup> (DE3), and Rosetta<sup>TM</sup> (DE3), ClearColi<sup>®</sup> BL21 (DE3), and SHuffle T7 (Table 3.1) (Doron). The BL21 (DE3) strain is a widely used, high expression strain that allows for expression of recombinant genes under the T7 promoter. The BL21 Star<sup>TM</sup> (DE3) strain is a derivative of the BL21 (DE3) strain with reduced levels of endogenous RNases resulting in more stable mRNA and enhanced protein expression. The Rosetta<sup>TM</sup> (DE3) strain is a variation on the BL21 (DE3) strain that supplies tRNAs that are not naturally expressed at high levels in *E. coli* to allow for increased production of eukaryotic proteins. The

ClearColi BL21 (DE3) is a genetically modified version of BL21 (DE3) with seven mutations affecting lipopolysaccharide (LPS) by deleting the oligosaccharide chain from the LPS and removing two of its six acyl chains; these modifications to LPS result in no endotoxin production by the strain. Endotoxins result in an innate immune response in humans that can lead to septic shock if LPS levels are too high (Sampath, 2018). Current methods of protein production with *E. coli* often require additional purification steps to remove endotoxins. Removing them at the source allows for a simplified purification train and a safer product. SHuffle T7 is an engineered *E. coli* K12 strain that promotes disulphide bond formation in the cytoplasm. It also constitutively expresses disulphide bond isomerase (DsbC). This should allow for improved expression of proteins with disulphide bonds.

Table 3.1 *E. coli* strains for Extract Preparation (Doron) (Lucigen, 2018)

<b><i>Bacterial Strain</i></b>	<b><i>Features</i></b>
<i>BL21 (DE3)</i>	Contains T7 polymerase upon IPTG induction; deficient of lon and omp-t proteases; suitable for expression of non-toxic genes
<i>BL21 Star (DE3)</i>	Contains T7 polymerase upon IPTG induction; contains a mutated RNase E to reduce RNase degradation and boost protein expression
<i>Rosetta (DE3)</i>	Contains T7 polymerase upon IPTG induction; supplies tRNA for the codons AUA, AGG, AGA, CUA, CCC, and GGA to enhance expression of eukaryotic proteins
<i>ClearColi (DE3)</i>	Contains T7 polymerase upon IPTG induction; contains a genetically modified lipopolysaccharide to eliminate endotoxins
<i>SHuffle T7</i>	Contains T7 polymerase upon IPTG induction; promotes disulphide bond formation in the cytoplasm and constitutively express DsbC

The cell extracts were used in CFPS reactions with the concentrated reaction mixture based on Kwon and Jewett (2015) and 6.1 nM (10 µg/mL) of the

pJL1 plasmid. The procedure used to set up these reactions can be found in Section 2.4.3. The reactions were analysed based on sfGFP production. The highest titres, an average of 506  $\mu\text{g/mL}$  sfGFP, were achieved with the Shuffle T7 extract, although the BL21 Star<sup>TM</sup> (DE3) extract and the ClearColi<sup>®</sup> BL21 (DE3) were also high producing, an average of 497  $\mu\text{g/mL}$  sfGFP and 492  $\mu\text{g/mL}$  sfGFP respectively (Figure 3.1). The BL21 (DE3) and Rosetta<sup>TM</sup> (DE3) extracts produced significantly less, 283  $\mu\text{g/mL}$  and 131  $\mu\text{g/mL}$  respectively. This suggests that mRNA stability plays a key role in protein production via CFPS, although significant protein production can still be achieved without the mutated RNase E found in BL21 Star<sup>TM</sup>. The inclusion of eukaryotic tRNAs has an adverse effect on protein production, as seen with the Rosetta<sup>TM</sup> (DE3) extract. However, this extract may be useful for the production of other proteins of eukaryotic origin in future. Overall, the varied sfGFP titres achieved with the different strains suggests that choosing the appropriate strain for the product of interest is important to achieving high titres. Due to the high performance of the BL21 Star<sup>TM</sup> (DE3) extract, it was used in the subsequent screening studies.

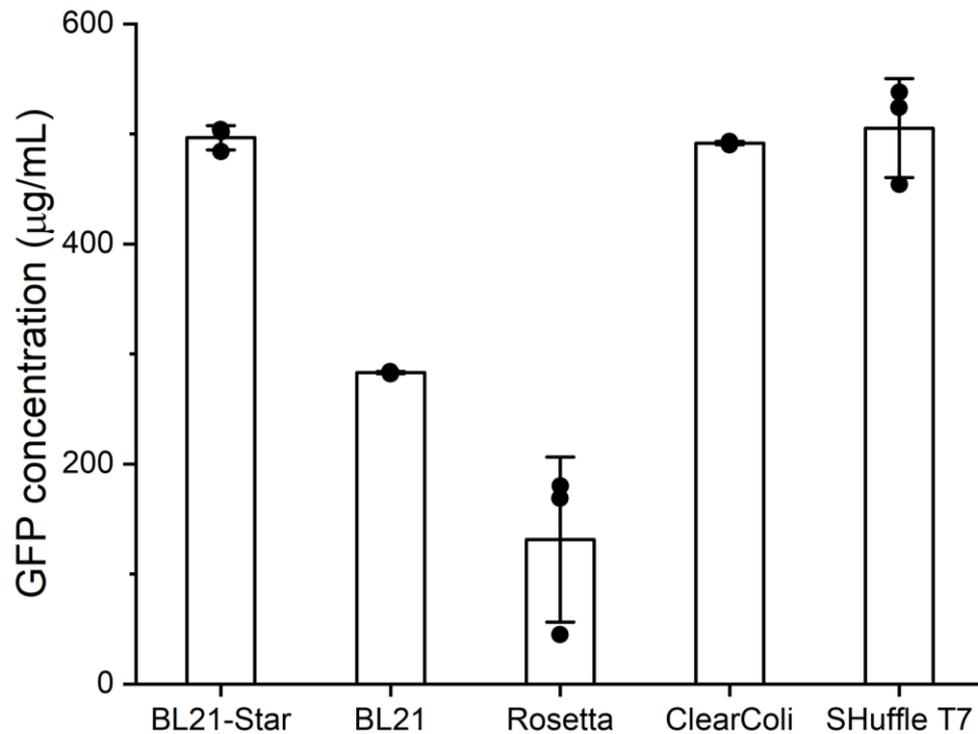


Figure 3.1 *E. coli* Extract Comparison

Extracts were prepared from five different strains and combined with concentrated reaction mixture based on Kwon et al. (2015) and the pJL1 plasmid. Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point. In typical reactions, the concentrated reaction mix was combined with 6.1 nM (10 µg/mL) pJL1 plasmid and 20% v/v cell extract and then incubated at 30°C for 4.0 hours.

### 3.3 Evaluating Complex and Minimal Concentrated Reaction Mixtures

Two concentrated reaction mixes, a complex mixture based on the work shown in Kwon and Jewett (2015) sometimes called the PANOx-SP system and a minimal mixture based on the work shown in Cai et al. (2015) referred to as the Cytomim system were prepared (Kwon and Jewett, 2015, Cai et al., 2015). For the CFPS reactions, these concentrated reaction mixes were combined with the non-induced BL21 Star™ (DE3) extract and 6.1 nM (10 µg/mL) of the pJL1 plasmid. The reactions were analysed based on sfGFP production. The complex mixture gave titres over three times greater than the minimal mix, an average of 497 µg/mL compared to an average of 146 µg/mL (Figure 3.2). This may be due to a depletion of energy sources in the minimal mix; the complex mixture contains more energy

sources, in particular phosphoenolpyruvate (PEP), coenzyme A (CoA), and nicotinamide adenine dinucleotide (NAD<sup>+</sup>), that may allow for prolonged ATP regeneration (Dopp et al., 2019b). The complex mixture also utilises nucleotide triphosphates instead of nucleotide monophosphates, as used in the minimal mix, which may allow for better ATP regeneration as well as higher rates of transcription and translation, especially when paired with additional *E. coli* tRNAs, which are also absent from the concentrated minimal mix. However, it is worth noting that CFPS reactions have been shown to be able to generate nucleotide triphosphates from nucleotide monophosphates in both crude cell lysate CFPS and the PURE system (Cai et al., 2015, Wang et al., 2019b). Based on these results, the complex concentrated reaction mixture was therefore used in subsequent screening studies.

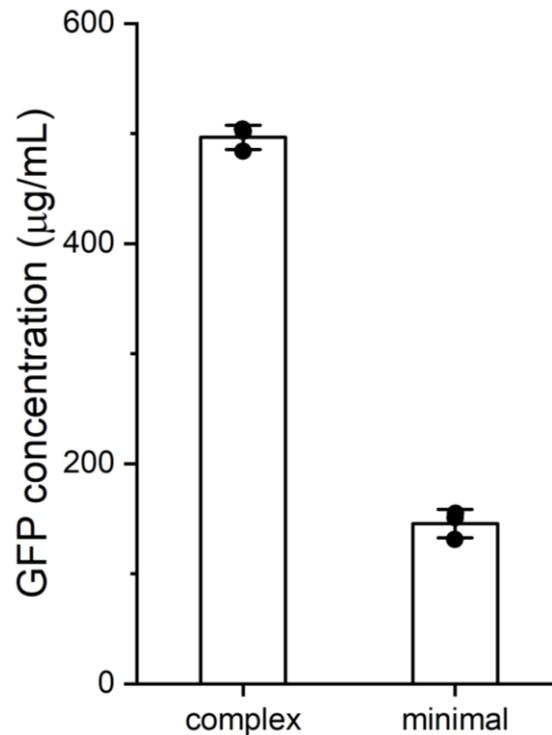


Figure 3.2 Concentrated Reaction Mixture Comparison

A complex concentrated reaction mixture based on the protocol in Kwon and Jewett (2015) and a minimal concentrated reaction mixture based on the protocol in Cai et al. (2015) were prepared and combined with the BL21 Star™ (DE3) extract and the pJL1 plasmid. Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point. In typical reactions, the concentrated reaction mix was combined with 6.1 nM (10 µg/mL) pJL1 plasmid and 20% v/v cell extract and then incubated at 30°C for 4.0 hours.

### **3.4 Examining Plasmids for In Vivo Expression and Plasmids Optimised for CFPS**

Before using the CFPS system to produce more difficult products, it was important to know whether the system was capable of expressing detectable amounts of protein when a plasmid that had previously been used for *in vivo* expression was utilised in the *in vitro* CFPS reaction or if modifications to the plasmid were critical for expression. Two plasmids were used: pJL1, a CFPS-optimised plasmid created by the Jewett Lab at Northwestern University for sfGFP production, and pET14b-GFP, a plasmid design for *in vivo* expression of GFP+ with a 6x histidine tag originally developed by Martin Warren's group at the University of Kent. In order to compare these two different versions of GFP, the

fluorescence readings were normalised based on the extinction coefficient and the quantum yield of each protein as described in Section 2.5.1. Both plasmids were used in CFPS reactions at a concentration of 3 nM where all other conditions were the same. The non-induced BL21 Star™ (DE3) extract and the complex concentrated reaction mixture based on the protocol in Kwon and Jewett (2015) were used. The reactions were analysed based on sfGFP and GFP+ production. The reactions with the pJL1 plasmid achieved over double the titre achieved with the reactions using the pET14b-GFP plasmid: an average of 227 µg/mL compared to an average of 106 µg/mL (Figure 3.3).

This demonstrates that detectable amounts of protein can be achieved with a traditional plasmid designed for *in vivo* production using the CFPS system which means optimisation is not necessary before using a plasmid in the CFPS system. This is not entirely unexpected as other groups have used plasmid backbones typically used *in vivo* in the past (Krinsky et al., 2016). However, traditional plasmids may not perform as well as a plasmid optimised for CFPS. Care should be taken, however, not to assume that this is generally true (Chizzolini et al., 2017, Iskakova et al., 2006, Jewett et al., 2016a, Jewett et al., 2016b). The value in first screening alternative plasmids comes also from the need to confirm the quality of the plasmids and their suitability for *in vitro* transcription and translation before undertaking the more extensive range of experiments which follow; it has been seen, for example, that plasmid purification may have a substantial effect on subsequent *in vitro* reactions (Strychalski and Romantseva, 2020).

The pJL1 plasmid was chosen for the subsequent experiments involving GFP for the sake of easier detection and easier plasmid preparation. The pJL1 plasmid expresses superfolder GFP (sfGFP), which folds more readily and is brighter than the GFP+ produced using the pET14b-GFP plasmid (Pedelacq et al.,

2006, Overkamp et al., 2013). This means that the fluorescence readings for sfGFP are generally stronger. In addition, the pJL1 plasmid is 2486 bp in size while the pET14b-GFP plasmid is 5389 bp in size. Using plasmid preparation kits, often times more pJL1 plasmid (~400-500 ng/ $\mu$ L, ~240-300 nM) than pET14b-GFP plasmid (~150-250 ng/ $\mu$ L, ~40-55 nM) was produced.

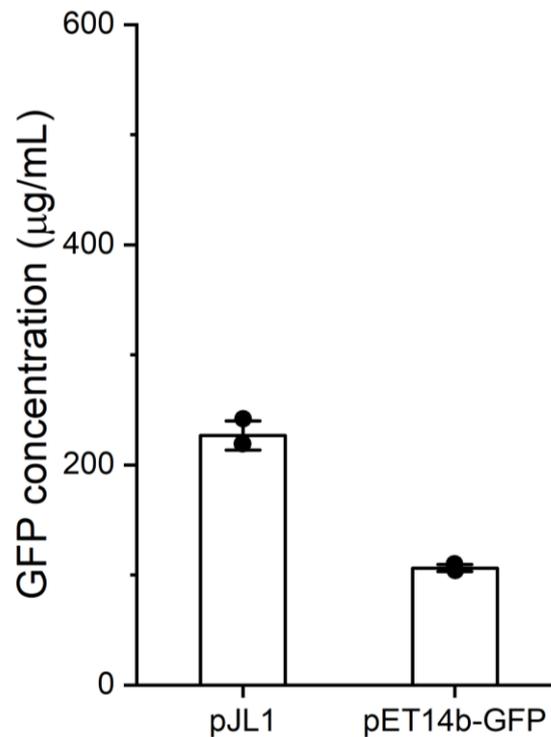


Figure 3.3 Plasmid Comparison

The pJL1 plasmid developed in the Jewett Lab at Northwestern University specifically for CFPS expression of sfGFP purchased from Addgene (unpublished) and the pET14b-GFP plasmid used for GFP expression in *E. coli* in the UCL Biochemical Engineering department were combined with the BL21 Star<sup>TM</sup> (DE3) extract and the complex concentrated reaction mixture based on the protocol in Kwon and Jewett (2015). Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point. In typical reactions, the concentrated reaction mix was combined with 20% v/v cell extract and then incubated at 30°C for 4.0 hours.

### 3.5 Comparison to Commercial *E. coli*-based CFPS Kit

The combination of extract, concentrated reaction mix, and plasmid that gave the highest average titre was the BL21 Star<sup>TM</sup> (DE3) extract, the complex concentrated reaction mixture based on the protocol in Kwon and Jewett (2015) and 6.1 nM (10 µg/mL) of the pJL1 plasmid. This CFPS platform was compared to

the commercial kit sold by ThermoFisher Scientific™, the Expressway™ Mini-Cell Free Expression System (2011). The reactions were analysed based on sfGFP production. This platform, which gave an average titre of 497 µg/mL, performed as well as the commercial kit, which gave an average titre of 493 µg/mL (Figure 3.4).

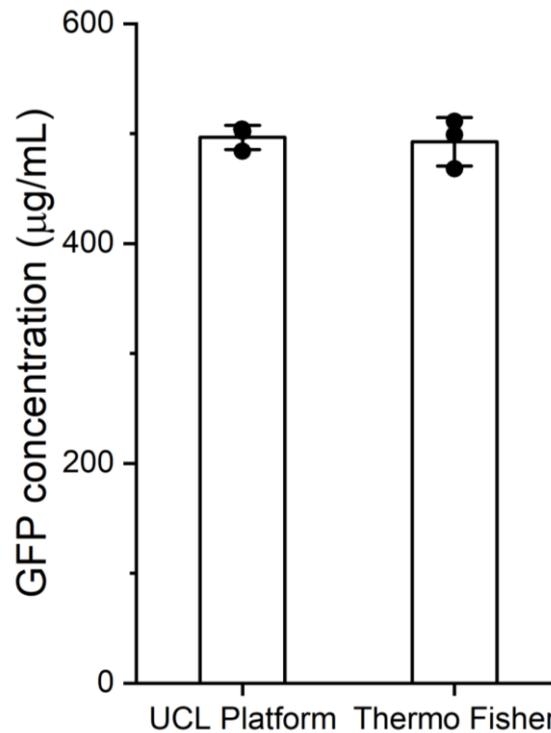


Figure 3.4 Comparison to Expressway™ Mini-Cell Free Expression System

The UCL platform which combines the BL21 Star™ (DE3) extract, the concentrated reaction mixture based on the protocol by Kwon and Jewett (2015) and the pJL1 plasmid was compared to the Expressway™ Mini-Cell Free Expression System sold by ThermoFisher Scientific™. Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point.

### 3.6 The Effect of IPTG-Induction in Cell Extract Preparation

While constitutive promoters like  $\sigma 70$  promoter can be used in CFPS, the gene of interest is under the T7 promoter in all the DNA plasmids used in this study (Failmezger et al., 2017, Silverman et al., 2019). The BL21 Star™ (DE3) strain, as with all DE3 strains examined in this study, can be induced with isopropyl  $\beta$ -D-1-thioglatopyranoside (IPTG) to induce the E. coli RNA polymerase (Figure 3.5) (Studier and Moffat, 1986). This allows for production of T7 RNA polymerase which results in the expression of the target gene. T7 RNA polymerase is an essential

component in the CFPS reaction; by increasing the concentration present in the crude cell extract, the amount of this expensive reagent in the reaction may be significantly decreased. An induced extract was prepared using the BL21 Star™ (DE3) strain and compared to the non-induced BL21 Star™ (DE3) strain, with and without additional exogenous T7 RNA polymerase in the reaction. The reactions were analysed based on sfGFP production as measured based on fluorescence (Figure 3.6) and further verified with a fluorescent-image of an SDS-PAGE (Figures 3.7). It was demonstrated that induced strains do not require additional T7 RNA polymerase, but non-induced strains do require additional T7 RNA polymerase. However, induced strains, with and without the additional T7 RNA polymerase, gave somewhat lower titres – an average of 357 µg/mL with additional T7 RNA polymerase and 379 µg/mL without additional T7 RNA polymerase – than the non-induced with additional T7 RNA polymerase – an average of 497 µg/mL. This suggests that induction may put some additional strain on the transcription and translation mechanisms in the CFPS reaction. It is important to be aware, however, of the substantial batch-to-batch variability which may be seen in extract production, as shown by the three additional batches of IPTG-induced extract in Figure 3.11 (Strychalski and Romantseva, 2020, Dopp et al., 2019a). While the batch-to-batch variation is not unreasonable by current standards in the art, it is still substantial relative to the titre difference. Consequently, for the pJL1 plasmid the difference in titre between the induced and non-induced extract with addition of T7 RNA polymerase is not statistically significant. The non-induced BL21 Star™ (DE3) extract was used in subsequent experiments with the pJL1 plasmid.

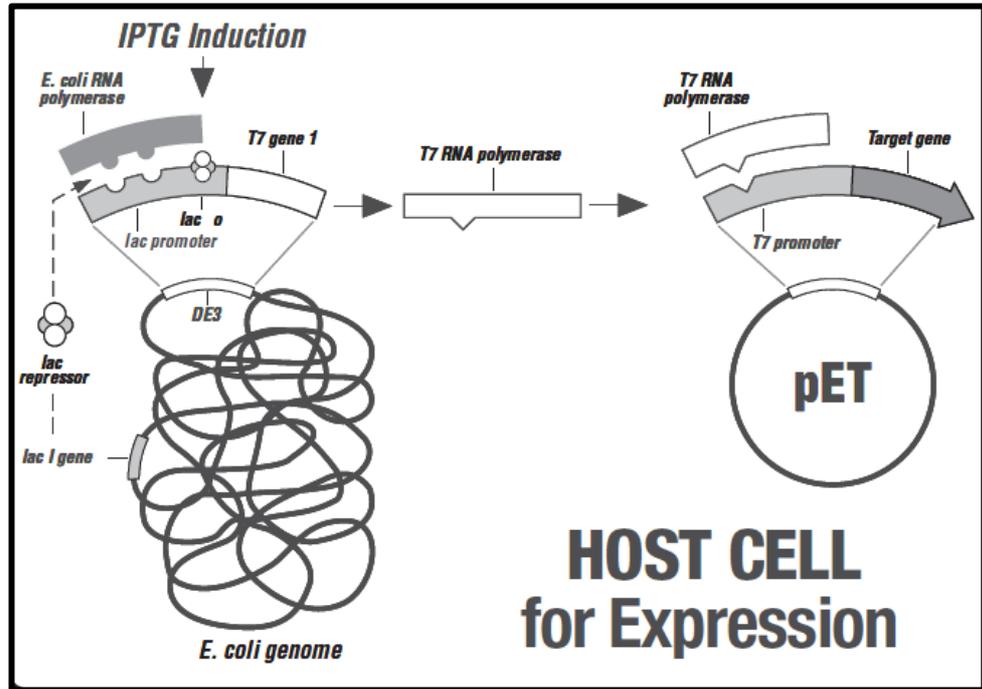


Figure 3.5 pET Plasmid Expression Under the T7 Promoter with IPTG Induction  
 Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) inhibits *E. coli* RNA polymerase resulting in the expression of T7 RNA polymerase for transcription of the target gene in a pET plasmid. Used with permission from (Novagen, 2018).

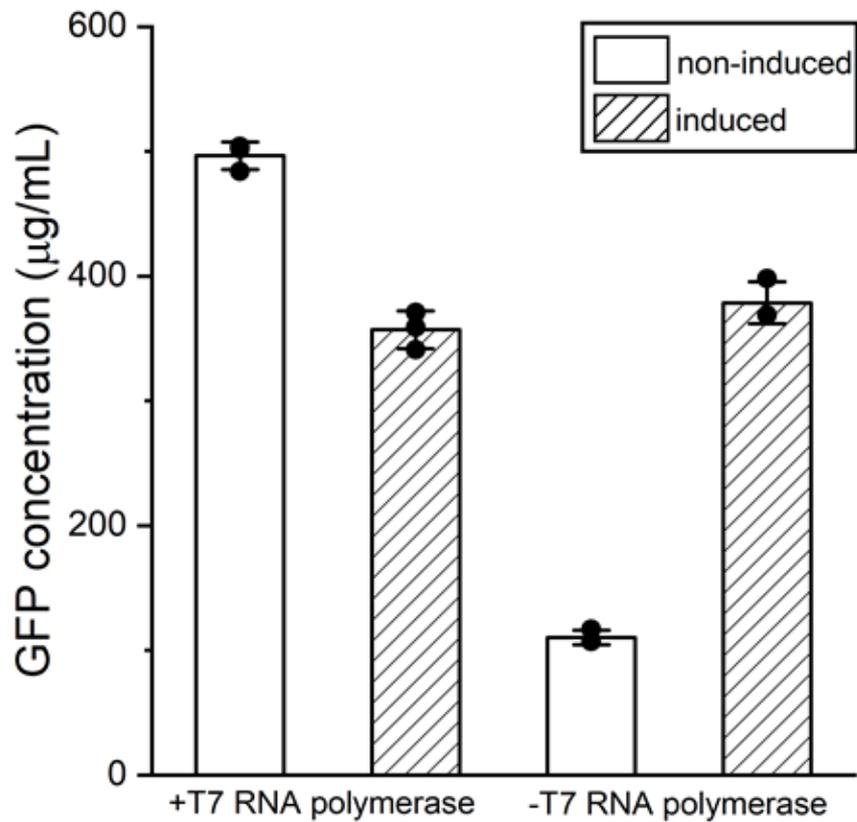


Figure 3.6 IPTG-induced and Non-induced BL21 Star™ (DE3) Extracts with and without Additional T7 RNA Polymerase

A non-induced and an induced BL21 Star™ (DE3) extract were combined with the pJL1 plasmid and the concentrated reaction mixture based on the protocol by Kwon and Jewett (2015) with and without additional T7 RNA polymerase. Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point.

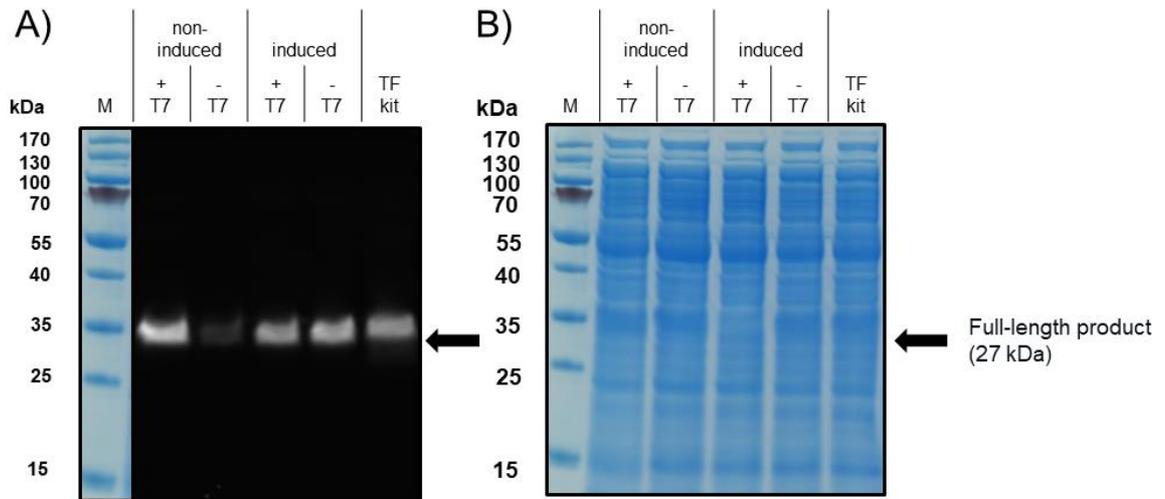


Figure 3.7 Fluorescent Image of SDS-PAGE of CFPS Reactions

An SDS-PAGE analysis of four CFPS reactions with either non-induced or induced BL21 Star (DE3) extract with or without additional T7 RNA polymerase (“+T7” or “-T7”) with the concentrated reaction mixture based on the protocol from Kwon and Jewett (2015) and the pJL1 plasmid and a fifth reaction with the pJL1 plasmid using the Expressway™ Mini-Cell Free Expression System from ThermoFisher Scientific™. CFPS reaction samples were diluted 1:3 in Milli-Q water, NuPAGE Sample Reducing Agent, and NuPAGE MES Sample Buffer. (A) The analysis was imaged under a blue fluorescent light (460 nm). (B) The analysis was imaged under white light.

### 3.7 Scale up of CFPS Reactions to 100 mL

Previously reactions were only performed at the 100  $\mu$ L scale. Here, 1 mL, 10 mL and 100 mL reactions were attempted. Reactions were performed in microcentrifuge tubes, deep well plates, baffled shake flasks, T-flasks, and DASbox bioreactors. An IPTG-induced BL21 Star™ (DE3) *E. coli*-based extract was used so that additional T7 RNA polymerase would not need to be added to the reactions. 6.1 nM (10  $\mu$ g/mL) of pJL1 plasmid was used to produce sfGFP. Titre analysis was based on fluorescence.

First, the current process was scaled up ten-fold in flower plates - deep-well plates shaped like hexagonal flowers which act as baffles to improve mixing in the wells. Although mixing is improved with deep-well flower plates, temperature transfer is more difficult due to the thick plastic layer of the bottom of the plate. The plates were pre-warmed to mitigate potential temperature issues. Reactions of 10

mL in volume were performed in 250 mL shake flasks and Sarstedt tissue culture T-75 flasks. Two different vessels were used for this volume of reaction in order to determine whether or not the geometry of the vessel had a significant impact on titre (no significant difference was observed). These vessels were fastened securely inside of a Kühner ISF1-X Climo-Shaker shaking incubator. The DASbox bioreactor system was used for 100 mL CFPS reactions. The use of bioreactors allowed for monitoring of pH and dissolved oxygen over the course of the reaction. Temperature was maintained at 30°C. The impeller speed was set to 500 rpm to ensure that the reaction was well-mixed.

The titres achieved across all five reactions vessels were very similar (Figure 3.8). Reactions in the microcentrifuge tubes yielded an average of 640 µg/mL. Those in flower plates, shake flasks, and T-75 flasks gave titres of 639 µg/mL, 627 µg/mL, and 627 µg/mL respectively. Only the bioreactors had a slightly lower titres, an average of 601 µg/mL. This suggests that the system scales linearly as expected, and that vessel geometry and size does not impact titre in a considerable way for the CFPS platform designed in this project. In the future, processes to express products of interest can be designed by first determining the appropriate conditions in microcentrifuge tubes and then scaling up to the desired volume depending on the amount of material needed.

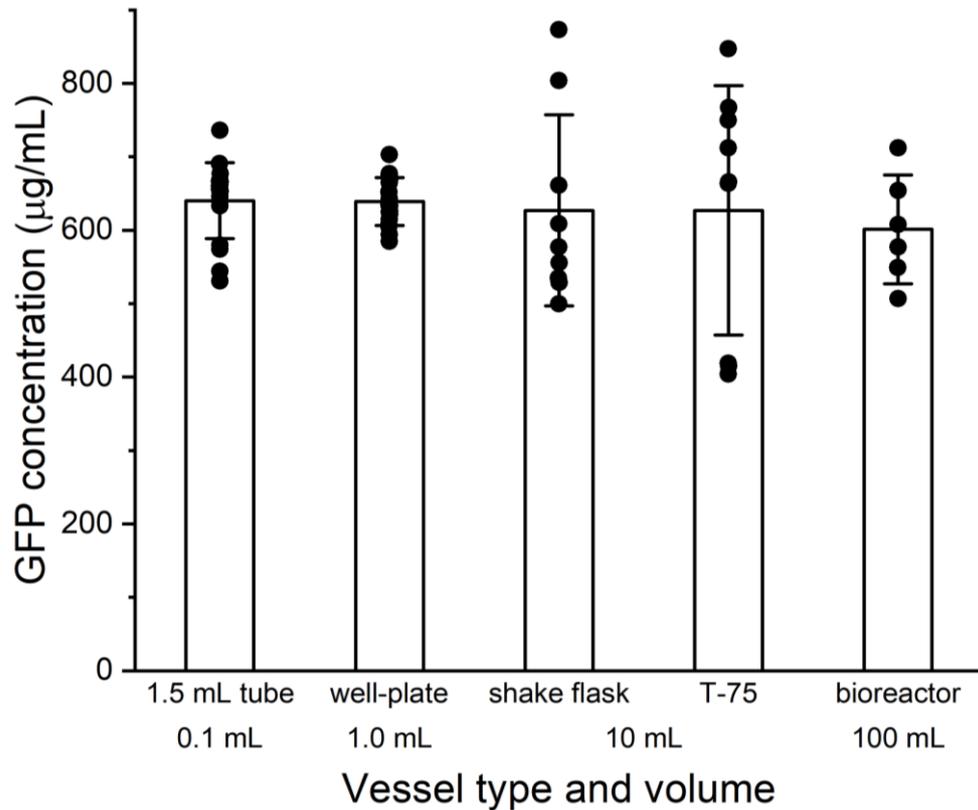


Figure 3.8 Scale up of sfGFP production over three orders of magnitude

An induced BL21 Star™ (DE3) extract was combined with the pJL1 plasmid and the concentrated reaction mixture based on the protocol by Kwon and Jewett (2015) without additional T7 RNA polymerase. Error bars represent plus or minus one standard deviation for  $n = 18$  technical replicates from 6 biological replicates for the 0.1 mL and 1 mL reactions, and  $n = 9$  technical replicates from 3 biological replicates for the 10 mL reactions and  $n = 6$  technical replicates from 2 biological replicates for the 100 mL reactions, each represented as a single data point. (More replicates were run for the smaller volume reactions because these results were collected as part of the Pilot Plant Week for UCL Biochemical Engineering undergraduate students and each of the six students on the team was given the opportunity to assemble and run a reaction.)



Figure 3.9 Images of sfGFP production

These images show the sfGFP that was produced in the vessels. The tubes in the left image contain the material harvest from all of the reaction vessels: 100 µL reactions in microcentrifuge tubes, 1 mL reactions in flower plates, 10 mL reactions in 250 mL shake flasks and T-75 flasks, and 100 mL in DASbox bioreactors. The centre image shows three T-75 flasks before harvest. The right image shows the DASbox bioreactor before harvest.

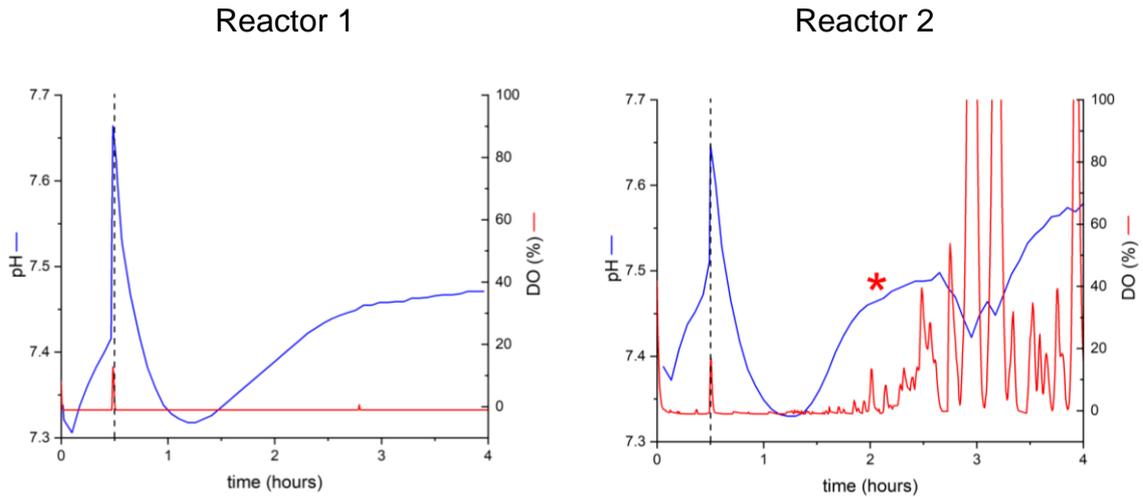


Figure 3.10 pH and DO levels of bioreactors

The diagrams show the pH of the entire reaction in blue and the DO levels in red for each vessel. The pH of the entire reaction began at around pH 7.3 when the first half of the reaction was put in the bioreactor. The level rose to pH 7.65 when the second part of the reaction was added after 30 minutes, shown with the dotted line. The DO levels in both reactors rose to ~10% with the addition of the second part of the reaction and then plummeted to 0% for the rest of the reaction. In the second vessel, one of the filters was fouled around 2h, shown with an asterisk (\*) and the DO levels proceeded to be erratic after that time.

In addition to demonstrating the scalability of the CFPS system, performing CFPS reactions in a bioreactor allowed for better monitoring of the conditions of the reaction. Over the course of the 4-hour reaction, the pH steadily declined from pH 7.6 to pH 7.2. CFPS reactions also consume a lot of oxygen. The DO levels of the reactors were at 0% for nearly the entire run. However, it does not appear that the reactions are oxygen limited, as the geometry and mixing of each of the platforms used was slightly different, but the titres were not dramatically different.

### **3.8 Examining IPTG-Induced BL21 Star<sup>TM</sup> (DE3) Cell Extract Batch-to-Batch Consistency**

Because the extract preparation protocol requires several steps that may ultimately impact the activity of the extract, each extract batch was tested by expressing sfGFP in the typical reaction conditions at the 100  $\mu$ L scale. The most commonly used extracts are the BL21 Star<sup>TM</sup> (DE3) non-induced and induced extracts. For both extracts, if the titre falls between 400  $\mu$ g/mL and 600  $\mu$ g/mL

sfGFP, the extract will continue to be used. Otherwise, it will be discarded, and a new batch of extract will be prepared. Figure 3.11 demonstrates a test of three BL21 Star™ (DE3) induced extracts prepared on 7 June 2019, 24 July 2019, and 11 March 2020 that achieved titres of 490 µg/mL, 428 µg/mL and 440 µg/mL respectively. Because the titres achieved with these three extracts fall in the 400-600 µg/mL they were all saved for future use.

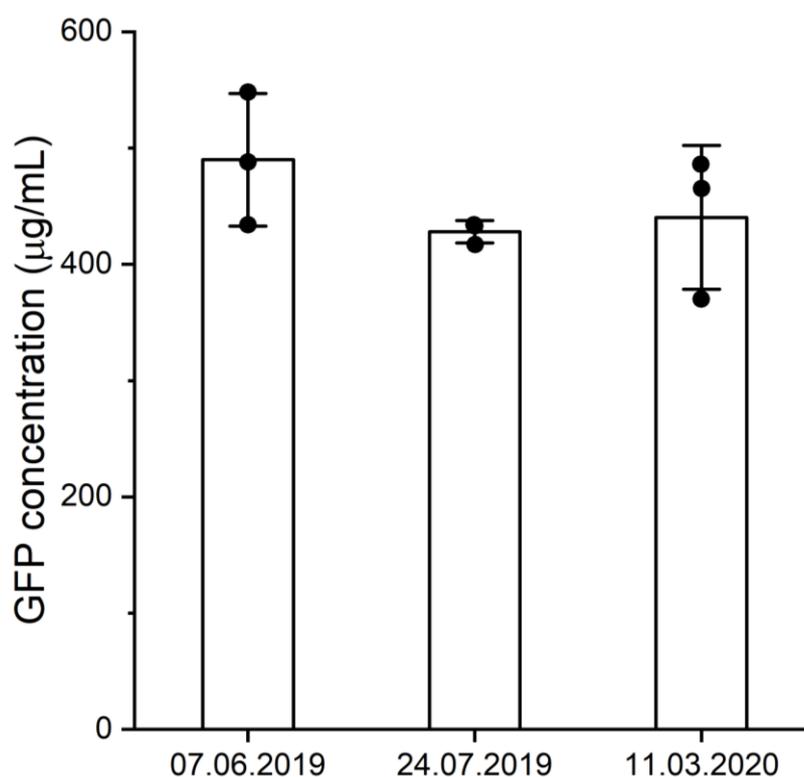


Figure 3.11 BL21 Star™ (DE3) Induced Extract Consistency Test

BL21 Star™ (DE3) induced extract prepared on three different dates was tested in a reaction with concentrated reaction mixture based on Kwon et al. (2015) and the pJL1 plasmid. Error bars represent plus or minus one standard deviation for  $n = 3$  biological replicates, each represented as a single data point. All extracts gave average titres in the 400-600 µg/mL range.

### 3.9 Conclusion

In this section, an in-house CFPS platform has been designed and demonstrated to scale up in a variety of reactor vessels. Extract strain, concentrated reaction mixture formulation, and plasmid design were examined. Higher titres were achieved with BL21 Star™ (DE3), ClearColi® BL21 (DE3), and SHuffle T7 extracts than with BL21 (DE3) and Rosetta™ (DE3) extracts. This

suggests that the more stable mRNA achieved with the BL21 Star™ (DE3) extract, the lack of endotoxin achieved with the ClearColi® BL21 (DE3) extract, and the constitutive expression of DsbC in the SHuffle T7 extract have a positive impact on protein production. For the concentrated reaction mixture, while the minimal mixture may be more cost-effective, the complex mixture results in higher titres, most likely because there are more potential energy sources in the complex mixture that allow the reactions to be sustained for longer. Using a plasmid design specifically for CFPS also boosts titre; however, detectable titres can be achieved with a typical plasmid for *in vivo* expression which means they can be used for proof of concept reactions. Additionally, scale up over three orders of magnitude is possible with this CFPS platform using vessels like deep-well flower plates, baffled shake flasks, T-75 flasks, and small-scale bioreactors. Extract activity has also been validated based on a simple sfGFP production test to minimise batch-to-batch variation as much as possible. In the future, small scale reactions in microcentrifuge tubes will be used to determine the reaction conditions for new products before the reaction volume is increased to generate the appropriate amount of material.

## **4. Designing a Process Development Strategy to Improve Titres of CFPS**

### **Reactions**

*A CFPS platform that performs as well as commercially available CFPS kits was developed in the previous chapter. Robust methods were needed to determine the appropriate reaction conditions for expression in the CFPS system so a process development strategy for E. coli-based CFPS reactions that can be completed in as little as 48 hours was designed. In the previous chapter, it was observed that the most dramatic increases in titre were due to the E. coli strain for the cell extract. Therefore, identifying a high-producing cell extract for the given product of interest was recommended as a first step. Next, the plasmid concentration, amount of extract, temperature, concentrated reaction mixture pH levels, and length of reaction were manipulated. The influence of these process parameters on titre was evaluated through multivariate data analysis (MVDA). The process parameters with the highest impact on titre were subsequently included in a design of experiments (DoE) enabling the optimum reaction conditions that maximise titre in a given design space to be determined. This proposed process development strategy resulted in superfolder green fluorescent protein (sfGFP) titres of 0.686 g/L, a 38% improvement on the existing operating conditions, and hepatitis B core antigen (HBcAg) titres of 0.386 g/L, a 190% improvement. Now that this process development strategy has been validated with two model products, production can be expanded to more complex products that are representative of personalised medicines.*

### **4.1 Introduction**

Although CFPS has many advantages and applications, little work has been done on process development for these reactions. Current process development strategies for a typical biopharmaceutical in a mammalian cell host involves lengthy

cell line development schemes that can take 60-90 days (Castan et al., 2018). While this process may be shorter for microbial hosts like *Escherichia coli* and *Pichia pastoris*, it is still cumbersome and requires cloning and selection from several strains. The fact that CFPS reactions can produce good titres in a matter of hours and do not necessarily require any cloning steps, as PCR products can be used, allows for the use of high-throughput methodologies to design a process development strategy (Schinn et al., 2016).

This was achieved by examining the impact of process parameters, in particular cell extract strain, on product titre. Certain *E. coli* strains that have been engineered to have more stable mRNA or more eukaryotic tRNAs may result in better translation and increased yields for certain products. Similarly, the impact of induction on product titre in IPTG-inducible BL21 Star™ (DE3) *E. coli* cells has been observed. Induction of the *E. coli* cells prior to extract preparation should result in an increased concentration of T7 RNA polymerase, an enzyme vital to transcription. Adding T7 RNA polymerase into the CFPS reaction is commonly found in the literature as T7 RNA polymerase is highly selective for its own promoter sequences and it has a transcription rate that is higher than endogenous *E. coli* RNA polymerase (Tabor, 1990).

Because CFPS reaction conditions are no longer constrained by the needs of maintaining metabolically active cells, CFPS reaction parameters can be extended to greater extremes which may aid in the synthesis of difficult-to-express proteins and self-assembly processes (e.g. virus-like particles) (Jin and Hong, 2018, Bundy et al., 2007, Sheng et al., 2017). By adding more of the potentially limiting components for example, plasmid DNA or cell extract, the reaction equilibrium may be shifted to increase product yield. However, as both plasmid and extract preparation are time-consuming and labour-intensive processes, it is critical

that the system not use more than what is required for either component. As the reactions are no longer limited by cell growth, the pH of the concentrated reaction mixture can be decreased or increased far beyond typical physiological levels. However, pushing the pH of the reaction too far may have an adverse effect on the molecules required for transcription and translation of the product (ribosomes, polymerases, enzymes, etc.) or could result in the precipitation of other reaction components. Because the polypeptide elongation rate in *E.coli* cells is enhanced at higher temperatures, increasing the reaction temperature should increase production, although too much of an increase may lead to protein degradation (Farewell and Neidhardt, 1998). Additionally, lower temperatures may be beneficial as CFPS reactions are considered less thermostable than their corresponding host cells because they are more dilute; lower temperatures may also aid in protein folding and prevent aggregation (Schumann and Ferreira, 2004). The length of the reaction should be long enough to allow for the expression of high titres of protein but should not be so long that inhibitors like inorganic phosphate saturate the system (Kim and Swartz, 1999).

The work presented here investigates the impact of the aforementioned process parameters on product titre. The *E. coli* strain used for the cell extract, two compositions of the concentrated reaction mix, and the plasmid selection were examined in the previous chapter. Here, multivariate data analysis (MVDA) was used to generate a model based on the titres resulting from reactions where the plasmid and cell extract concentration, pH of the concentrated reaction mix, reaction temperature, and length of reaction were manipulated individually. (A comparison of two concentrated reaction mixtures was chosen rather than an investigation of the individual components of the concentrated reaction mixtures because this has already been examined in some depth elsewhere (Cai et al.,

2015, Dopp et al., 2019b).) By performing a multilinear regression (MLR), combinations of parameters were predicted that result in the highest titres within the robust operating space defined. The process parameters with the largest influence on titre were further evaluated through a design of experiments (DoE) approach enabling the operating conditions to a significant increase in product titre to be identified. Several other groups have used DoE previously to examine multiple parameters at once while using a minimal amount of reaction material in order to optimise parts of the CFPS system including extract preparation, chaperone and salt concentrations for expression of proteins with disulphide bonds, and the ratio of heavy chain expressing plasmid to light chain expressing plasmid for antibody expression (Dopp and Reuel, 2018, Goerke and Swartz, 2008, Zawada et al., 2011, Yin et al., 2012). Here DoE was used for titre maximisation in a given design space, rather than optimisation, to demonstrate how this process development strategy might be used with two different proteins.

#### ***4.2 Examining a Variety of Process Parameters Using sfGFP***

The influence of varying the following process parameters on sfGFP titres were examined: plasmid concentration, amount of extract, temperatures, pH of the concentrated reaction mix, and lengths (Figure 4.1). Each parameter was initially examined in isolation. The initial values were chosen based off the recommendations made in the ThermoFisher Scientific™, the Expressway™ Mini-Cell Free Expression System handbook (2011). For a standard reaction, the following conditions were used: 6.1 nM (10 µg/mL) plasmid, 20% v/v extract, 30°C, a concentrated reaction mixture at pH 6.8, and 4 hr long reaction. The BL21 Star™ (DE3) extract without IPTG induction, the complex concentrated reaction mix, and the pJL1-sfGFP plasmid were used.

Commercial CFPS kits recommend 10 µg/mL plasmid and 20% v/v extract (2011). However, the processes for plasmid and extract preparation are laborious, time-consuming, and expensive. Therefore, minimising the amount used in each reaction would enable more experimental conditions to be evaluated faster and more economically. However, if reactions with higher plasmid or extract concentrations resulted in significantly higher titres this would be beneficial knowledge for process development. With that in mind, plasmid concentrations from 0.61 nM (1 µg/mL) to 30.4 nM (50 µg/mL) and amounts of extract from 5% v/v to 35% v/v were examined. It was observed that increasing the concentration of plasmid and the amount of extract improved sfGFP titres, but that the increase in titre eventually plateaus (Figure 4.1A and B). Reactions with amounts of above 20% v/v extract and concentrations of plasmid above 6.1 nM (10 µg/mL) resulted in very similar titres (~450 µg/mL sfGFP). It is likely that other resources (polymerases, amino acids, nucleotides, etc.) are depleted and plasmid or extract is no longer the limiting reagent.

Commercial kit suppliers recommend reaction temperatures between 30°C and 37°C (2011). This range was expanded to test reactions at temperatures from 15°C to 40°C. Titres peaked with a reaction temperature of 35°C (Figure 4.1C). The data suggest that a temperature between 32°C and 35°C is the ideal temperature for production of sfGFP via CFPS. Although 32-35°C may maximise sfGFP production, other products may require higher or lower temperatures. Lower temperatures might be preferable for more complex molecules with solubility issues as this tends to reduce the formation of inclusion bodies (de Groot and Ventura, 2006).

Previous studies have indicated that pH is one of the most critical process parameters in CFPS reactions (Caschera and Noireaux, 2015a, Caschera and

Noireaux, 2015b). In sfGFP production, this was also observed to be true; as the pH of the concentrated reaction mixture decreased, the product concentration increased (Figure 4.1D). Titres of over 700  $\mu\text{g/mL}$  were achieved with a concentrated reaction mixture of pH 5.5. This is likely because the other components in the CFPS system are basic in nature, in particular the concentrated solutions of amino acids that are added separately to the reaction, which must be kept at pH 12 in order to remain soluble (Dopp et al., 2019a). By using the concentrated reaction mixture to decrease the overall pH, the reaction as a whole was closer to a more neutral pH which may be ideal for the transcription and translation machinery in the extract. For example, the pH of the entire reaction that had had a concentrated reaction mixture of pH 5.5 was measured after it reached completion (4 hours) and its pH level was 7.2. Reactions at even lower pH values might be achieved in the future by preparing the extract with an acidic buffer and minimising the addition of base required to keep the amino acids in solution. Also, it is important to note in interpreting these results that while sfGFP is a pH sensitive protein, it has been demonstrated to display negligible differences in fluorescence intensity in the pH 5.3-9.4 range and the reaction is diluted 10-fold with 20% extract before measurement which should minimise the effects of the slight differences in overall pH between the reactions (Roberts et al., 2016, Stepanenko et al., 2014).

Reaction length is highly variable amongst previous studies: batch reactions from 2 hours to 24 hours have been examined (Kim and Swartz, 1999, Stark et al., 2018). In my own studies, I observed a visible green tint to the CFPS reactions after only 0.5 hours of incubation; therefore, reaction lengths from 0.5 hours to 22 hours were examined. In observing the length of the reactions, titres stabilised after 4 hours (Figure 4.1E). However, sfGFP is known to fold efficiently with good folding kinetics (Pedelacq et al., 2006). Other products with known assembly issues may

require longer reaction times. Alternatively, certain amino acids and nucleotides may be depleted after 4 hours of reaction. To replenish these reagents, a concentrated solution of amino acids and nucleotides could be fed into the reaction or continuous reactions could be used instead of the batch method employed here (Kim and Choi, 1996).

#### **4.3 Further Analysis Using Hepatitis B Core Antigen**

In order to understand the interaction between process parameters in CFPS reactions and product type, the study was extended to self-assembling hepatitis B core antigen (HBcAg) virus-like particles (VLPs). First, HBcAg production was observed in CFPS reactions with the non-induced BL21 (DE3), BL21 Star™ (DE3), and Rosetta™ (DE3) and the induced BL21 Star™ (DE3) cell extracts (Figure 4.2A). An induced extract was required for the production of significant titres of HBcAg. In fact, when non-induced and induced BL21 Star™ (DE3) extracts were used both with and without additional T7 RNA polymerase, two very different outcomes were observed (Figure 4.2B). Little to no expression was observed with the non-induced extract regardless of whether or not additional T7 RNA polymerase was used in the reaction. In reactions with the induced extract, HBcAg expression was easily detectable – again, whether or not additional T7 RNA polymerase was used in the reaction. Induced extracts should have a higher concentration of T7 RNA polymerase than non-induced extracts and would contain residual IPTG that would be found within the homogenised cell cytosol. This may be critical because the HBcAg gene is expressed from a pETDuet-1 plasmid and this plasmid contains the lac operator and the lac repressor gene (*lacI*) which normally inhibits transcription of the gene of interest. However, the inhibition is relieved when IPTG binds to LacI, and the gene of interest can be expressed. Alternatively, induction may result in other changes to the cellular components

(ribosomes, elongation factors, initiation factors, release factors, etc.) that may inherently improve protein expression, but that will depend upon the rate limiting step for a given protein/plasmid. Subsequent process parameter analysis was based on the results achieved using the IPTG-induced BL21 Star<sup>TM</sup> extract. The purified particles were observed under TEM and demonstrated self-assembly (Figure 4.2C).

The same five process parameters examined previously were investigated with the HBcAg plasmid, although the ranges were adjusted slightly based on the effects seen with sfGFP (Figure 4.1). As before, each parameter was examined in isolation. For a standard reaction, the following conditions were used: 2.6 nM (10 µg/mL) plasmid, 20% v/v extract, 30°C, a concentrated reaction mixture at pH 6.0, and 4 hr long reaction. The IPTG-induced BL21 Star<sup>TM</sup> (DE3) extract, the complex concentrated reaction mix, and the petDuet-1 plasmid with the HBcAg gene were used.

In some cases, similar trends to those seen with the sfGFP production were observed. Amount of extract from 5% v/v to 30% v/v was examined. Titres were very similar (~175 µg/mL) when amount of extract was above 15% v/v for monomeric HBcAg (Figure 4.1B). As observed with sfGFP, it is likely that other resources are exhausted, and extract is no longer the limiting reagent. Reactions were performed at temperatures from 20°C to 35°C. Much like sfGFP titres, HBcAg titres peaked with a reaction temperature between 32°C and 35°C (Figure 4.1C). Reaction lengths from 0.5 hours to 24 hours were examined. As with sfGFP, HBcAg titres stabilise after 4 hours (Figure 4.1E). However, as the length of the reaction increases, the variability in titre becomes much greater. Therefore, when possible shorter reaction times are recommended.

For both products, the highest titres are achieved when a plasmid concentration of over 5 nM (10  $\mu\text{g}/\text{mL}$  for sfGFP, 20  $\mu\text{g}/\text{mL}$  for HBcAg) is used; the product concentrations plateau or decrease at higher plasmid concentrations (Figure 4.1A). A few research groups have examined plasmid concentration with a single product and determined that increasing the plasmid concentration can boost product concentration to a certain point, but other species involved in transcription and translation need to be replenished after that, namely tRNAs and T7 RNA polymerase (Doerr et al., 2019, Hong et al., 2015, Nagaraj et al., 2017). It would seem that 5 nM of plasmid, regardless of the product, is the maximum amount the CFPS system in this project can accommodate before other components must be manipulated to increase product concentration. Also, aside from the choice of using an induced extract over a non-induced extract, plasmid concentration was the process parameter which had the greatest impact on monomeric HBcAg titre, which may suggest the typical amount of plasmid used in the reaction (2.6 nM) is limiting.

A different trend for pH of the concentrated reaction mixture than what was seen with sfGFP production was observed. HBcAg titres were not greatly affected by a change in pH of the concentrated reaction mixture (Figure 4.1D). The concentrated reaction mixture was prepared at pH values from pH 5.5 to pH 7.5. For HBcAg, the titres all lie within 100  $\mu\text{g}/\text{mL}$  of each other. This might indicate that reactions expressing HBcAg are limited by a component that is not particularly pH sensitive.

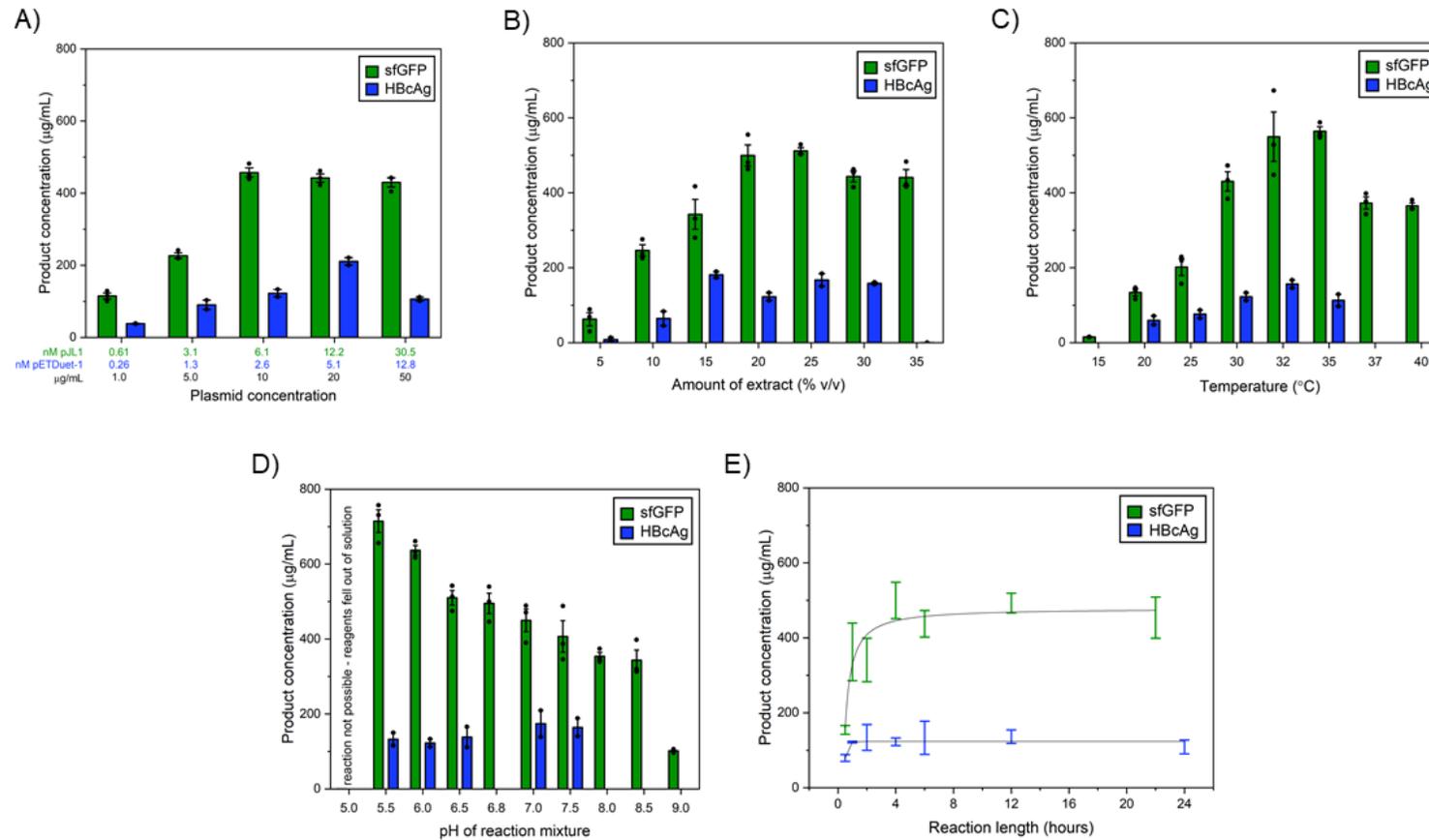


Figure 4.1 Process Parameter Effects on Product Titre

Titres for sfGFP are shown in green and titres for HBcAg are shown in blue. (A) Product concentration increases with increased plasmid concentration until 6.1 nM (10 µg/mL) for sfGFP and 5.1 nM (20 µg/mL) for HBcAg. (B) Likewise, product concentrations increase with increased amount of extract until 20% v/v for sfGFP and 15% v/v for HBcAg. (C) The highest product concentrations are seen at temperatures between 30°C and 35°C for both products. (D) Product concentration increases with decreasing pH of the concentrated reaction mix for sfGFP and product concentration is not significantly affected by pH for HBcAg; pH 5.0 could not be achieved due to precipitation of the concentrated reaction mix components. (E) The maximum product concentration is achieved after 4 hours of reaction. For (A) through (D), error bars represent plus or minus one standard error for n = 3 biological replicates for sfGFP and n = 2 biological replicates for HBcAg, each represented as a single data point. For (E) error bars represent plus or minus one standard error for n = 3 biological replicates for sfGFP and n = 2 biological replicates for HBcAg.

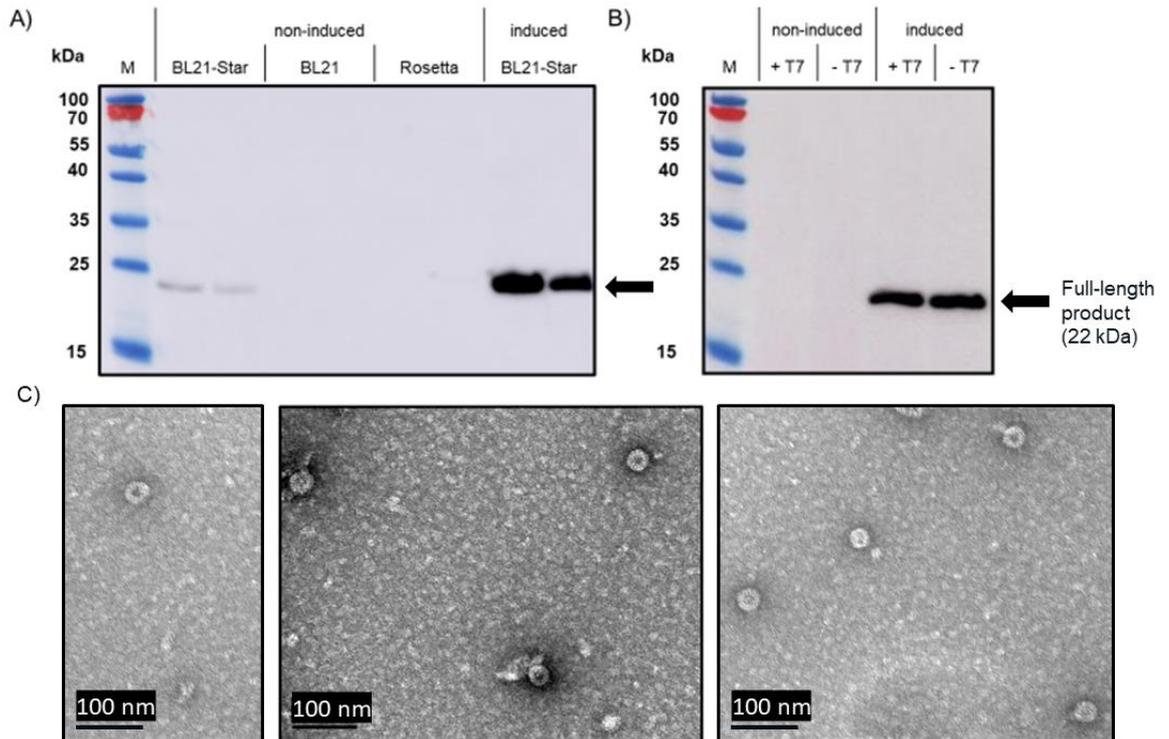


Figure 4.2 Extract Strain Effects on HBcAg Monomer Titre

(A) Induced BL21 Star<sup>TM</sup> extract results in a higher product concentration of HBcAg monomer compared to other non-induced extracts. For each extract, the left lane is the CFPS reaction and the right lane is the soluble fraction from that reaction. (B) No production of HBcAg is observed in non-induced BL21 Star<sup>TM</sup> (DE3) extract regardless of whether or not additional T7 RNA polymerase is added to the reaction; production is observed with and without additional T7 RNA polymerase in reactions with an IPTG-induced BL21 Star<sup>TM</sup> (DE3) extract. The reactions run in (B) are separate from the ones run in (A) and the IPTG-induced BL21 Star<sup>TM</sup> (DE3) extract used in (B) is a different batch than the one used in (A) therefore representing biological replicates. (C) Assembled HBcAg VLPs imaged under TEM from the reaction with the following conditions: 2.6 nM (10  $\mu$ g/mL) plasmid, 20% v/v extract, concentrated reaction mix pH 6.0, 30°C, and 4 hours. These TEM images were taken by the author under the supervision of Mark Turmaine at UCL.

#### ***4.4 Multivariate Data Analysis to Maximise Product Titre***

It was difficult to quantify the influence of each variable on product concentration due to the complex interactions between all process parameters. Therefore, multivariate data analysis (MVDA) was selected to evaluate the screening design based on its proven ability within the biopharmaceutical sector to leverage useful information from complex data sets and uncover useful correlations that are not always obvious from univariate analysis (Goldrick et al., 2020). The DoE methodology implemented is a systematic approach enabling the relationship between process operation and process output to be determined while reducing the required number of experiments to understand these key relationships. The face-centred composite (FCC) design of experiment was selected as it is the most appropriate design when factors investigated cannot be extended beyond the factorial points which was the case in this experiment. It also enables linear, interactive and quadratic terms to be evaluated as it contains centre points in addition to identifying the process conditions to maximise product concentration.

To assess the relative importance of each variable and their ability to predict the product concentration, four different types of MLR models were generated. These included linear, quadratic, interactions and squared relationships. The linear models considered an intercept and a linear term for each predictor. The squared model additionally accounted for squared terms. The interaction model considered the intercept term, linear relationships and all product pairs of distinct predictors. The quadratic model was similar to the interaction model and additionally accounted for squared terms of each predictor.

The relative importance of each variable was assessed by initially building each of these MLR models using all the predictor variables (Time, Plas, Ext, pH

and Temp) and then building a separate model with one of these predictor variables removed. The difference in the prediction ability of each of these models enables the relative importance of each variable to be defined. This systematic approach was performed for the four different MLR models (linear, quadratic, polynomial and interactions). Evaluating the difference in the root mean square error between these models enabled the relative contribution of each variable on titre to be quantified. An example of a quadratic model to predict the titres for sfGFP is shown in Figure 4.3A while the one for HBcAg is shown in Figure 4.3B. The MLR inputs used all predictor variables and were converted to coded factors to enable easier comparison of the coefficients shown in Figure 4.3C and 4.3D. The exact experiments used for Figure 4.3A are shown in Table 4.1 and those for Figure 4.3B are shown in Table 4.2. The relative importance of each input variable for this MLR model is determined by the magnitude of the coefficient with sign indicating where there is a positive or negative relationship between the input variable and product concentration. Within the univariate analysis, the coefficient of determination,  $R^2$ , was 0.78 for sfGFP and 0.70 for HBcAg. Equations for the MLR models shown in Figure 4.3 A-D of each product are shown below:

$$\begin{aligned} \text{sfGFP concentration} &\sim 0.45 + 0.63\text{Time} + 0.92\text{Plas} + 0.33\text{Ext} - 0.54\text{pH} \\ &\quad + 0.28\text{Temp} - 0.13\text{Time}^2 - 0.17\text{Plas}^2 - 0.09\text{Ext}^2 + 0.06\text{pH}^2 \\ &\quad - 0.12\text{Temp}^2 \end{aligned}$$

$$\begin{aligned} \text{HBcAg concentration} &\sim 0.65 + 0.34\text{Time} + 1.40\text{Plas} + 0.38\text{Ext} + 0.16\text{pH} \\ &\quad + 0.12\text{Temp} - 0.12\text{Time}^2 - 0.36\text{Plas}^2 - 0.10\text{Ext}^2 - 0.09\text{Temp}^2 \end{aligned}$$

A summary of the averaged contributions of each variable on product concentration calculated as previously described is shown in Figure 4.3E and 4.3F for the sfGFP and HBcAg, respectively. The pH of the concentrated reaction mixture and temperature were shown to have the largest influence on the

concentration of sfGFP and the plasmid concentration, temperature and amount of extract were found to have the largest influence on the concentration of HBcAg.

To validate the importance of the interactions of these parameters, a face-centred response surface DoE was performed varying these selected variables. For the sfGFP, the pH was varied between 5.5-6.5 and the temperature between 32°C and 36°C. The titres generated by this DoE resulted in similarly high sfGFP concentrations to those shown in Figure 4 and validated that low pH (5.5) and higher temperatures (34-36°C) resulted in maximal titres. These new conditions gave a titre of 686 µg/mL, a 38% increase from the typical reaction conditions (pH 6.8, 30°C, 6.1 nM (10 µg/mL) plasmid, 20% v/v non-induced BL21 Star™ (DE3) cell extract, and 4 hours) which resulted in titres of 497 µg/mL. The design of this set of experiments and the resulting contour plot based on interpolating the experimental product concentration between the experimental pH and temperature ranges can be seen in Figure 4.4. The contour plots shown are generated by interpolating the experimental product concentration between the experimental pH and temperature ranges investigated. For the HBcAg, the three variables manipulated by the DoE were the plasmid concentration (3.8-11.5 nM (15-45 µg/mL)), amount of extract (15-25% v/v) and temperature (32-36°C). The experiments generated by this DoE resulted in significantly higher titres than previous experiments with the maximum found at a temperature of 32°C, a plasmid concentration of 45 µg/mL and an extract concentration of 25% v/v. This titre was almost triple the previous highest titre and demonstrates the importance of performing such a titre improvement exercise. A titre of 386 µg/mL was observed, a 190% increase in titre from 133 µg/mL, which is achieved with the typical reaction conditions (pH 6.0, 30°C, 2.6 nM (10 µg/mL) plasmid, 20% v/v IPTG-induced BL21 Star™ (DE3) cell extract, and 4 hours). Another group used a modified A19 extract

and 12 nM plasmid (where the gene of interest was in a pET24a backbone) to generate similar titres (Bundy et al., 2007). The FCC DoE and results can be seen in Figure 4.5.

The maximum titre for both products is located at the edge of the design space suggesting titre could be further improved and optimised by widening the experimental design space considered in this work. However, it is not the intention to optimise titre in this work, merely to demonstrate the advantages of a systematic process development approach, as summarised in the following section, and to maximise titre within a given design space. It is also important to mention that expanding the design space could have other unintended consequences. For example, the optimum titre of HBcAg expression may be achieved with increased amounts of plasmid and extract, but significantly increasing the presence of these components in the reaction also increases the cost of the reaction.

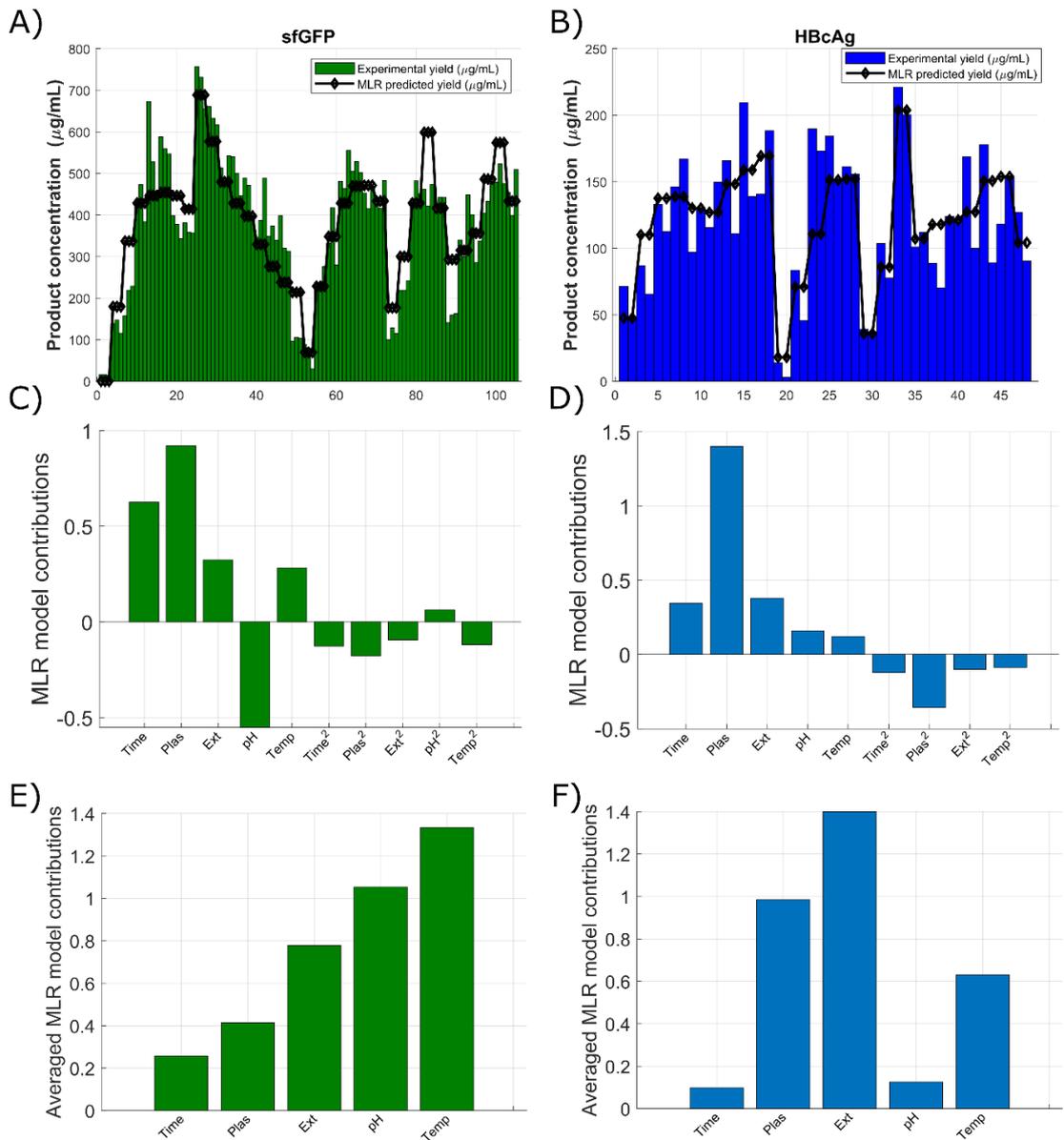


Figure 4.3 Multivariate Data Analysis of Process Parameters

The multilinear regression (MLR) models for sfGFP and HBcAg titres respectively are shown in (A) and (B) where each bar represents the product concentration from an experimental run (the experimental runs are the same ones shown in Figure 4.1 and detailed in Table 4.1 and Table 4.2). The contributions of each parameter to titre is shown in (C) and (D) where “Time” is the length of the reaction “Plas” is the plasmid concentration, “Ext” is the amount of extract in the reaction, “pH” is the pH of the concentrated reaction mix, and “Temp” is the temperature of the reaction. (E) and (F) show these contributions as normalised values.

Table 4.1 Experimental Conditions for Reactions Producing sfGFP Used for MVDA in Figure 4.3A

Experiment	reaction length (hours)	plasmid concentration ( $\mu\text{g}/\text{mL}$ )	plasmid concentration (nM)	extract (% v/v)	pH of concentrated reaction mixture	Temperature ( $^{\circ}\text{C}$ )	Product concentration ( $\mu\text{g}/\text{mL}$ )
1-24	4	10	6.1	20	6.8	15-40	16-673
25-51	4	10	6.1	20	5.5-9.0	30	96-757
52-72	4	10	6.1	5-35	6.8	30	30-555
73-87	4	1-50	0.61-30.5	20	6.8	30	100-482
88-105	0.5-22	10	6.1	20	6.8	30	141-523

Table 4.2 Experimental Conditions for Reactions Producing HBcAg Used for MVDA in Figure 4.3B

Experiment	reaction length (hours)	plasmid concentration ( $\mu\text{g}/\text{mL}$ )	plasmid concentration (nM)	extract (% v/v)	pH of concentrated reaction mixture	Temperature ( $^{\circ}\text{C}$ )	Product concentration ( $\mu\text{g}/\text{mL}$ )
1-10	4	10	2.6	20	6.0	20-35	48-167
11-18	4	10	2.6	20	5.5-7.5	30	111-209
19-28	4	10	2.6	5-30	6.0	30	3-190
29-36	4	1-50	0.26-12.8	20	6.0	30	37-221
37-48	0.5-24	10	2.6	20	6.0	30	70-178

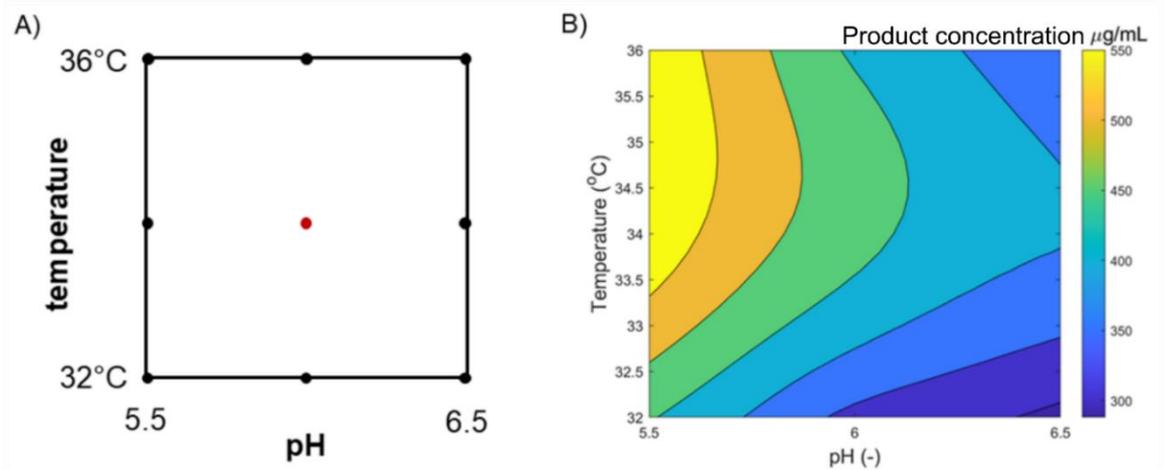


Figure 4.4 Design of Experiments for sfGFP yields

The experimental conditions for this face-centred central composite design are represented in (A) where each of the triplicate experiments performed is shown as a single point (9 reactions). The red dot in the centre represents the two centre points. The contour plot based off of the average yields generated through the DoE shown in (A) are shown in (B) where the interpolated product concentration in  $\mu\text{g/mL}$  is indicated by colour according to the gradient on the left-hand side.

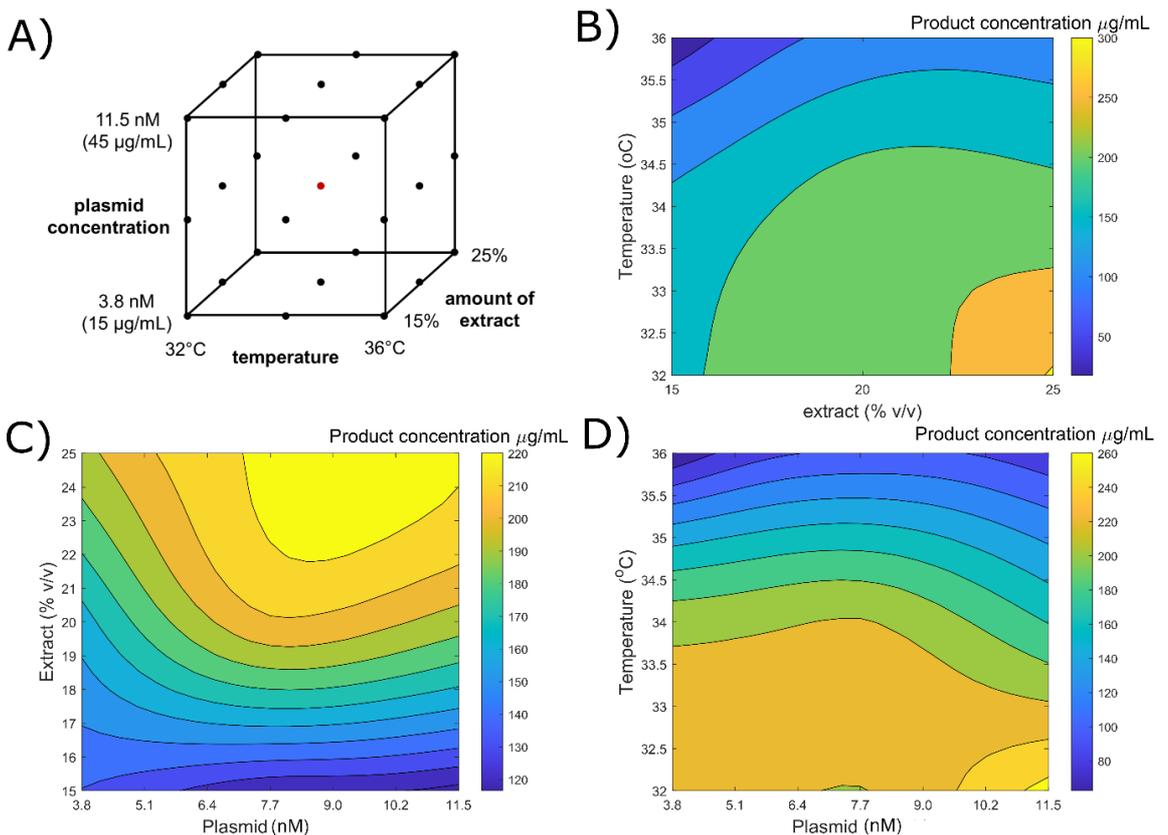


Figure 4.5 Design of Experiments for HBcAg yields

The experimental conditions for this face-centred central composite design are represented in (A) where each of the duplicate experiments performed is shown as a single point (27 total reactions). The red dot in the centre represents the two centre points. The influence of temperature, plasmid concentration, and amount of extract from the DoE in (A) are shown in (B), (C), and (D) where the interpolated product concentration in  $\mu\text{g/mL}$  is indicated by colour according to the gradient on the left-hand side. (B) shows the impact of temperature and amount of extract. (C) shows the impact of temperature and plasmid concentration. (D) shows the impact of the amount of extract and the plasmid concentration.

#### **4.5 Recommended Process Development Method**

Based on the findings in the previous sections, a recommended process development strategy for the maximisation of titre in CFPS reactions has been created (Figure 4.6). The appropriate cell extract must be determined first before considering any other parameters. The cell extract has the most significant impact on product titre (see Figure 3.1 in Chapter 3). While the use of the BL21 Star™ (DE3) extract resulted in high titres for the two products examined, other groups have used a variety of other extracts – including BL21 (DE3) and Rosetta (DE3) – to successfully express protein (Chumpolkulwong et al., 2004, Kim et al., 2006b). In addition, depending on the protein of interest, there are a number of other strains that might have properties that would be beneficial for the protein, for example, the use of amber-less strains when producing proteins with non-standard amino acids. Using a smaller plasmid optimised for CFPS is also recommended, however, there are several applications where maintaining a plasmid that could be used for more both CFPS and *in vivo* cultivation would be desirable. Also, optimisation of the concentrated reaction mixture for a particular product was not explored in this work, but based on previous examples, may also be beneficial for significantly increasing product titres (Cai et al., 2015, Dopp et al., 2019b, Kai et al., 2013, Pederson et al., 2011). Next, the influence of process parameters, changed in isolation, on titre and the application of MVDA to quantify the impact of those parameters should be evaluated. Subsequently the parameters with the highest influence on titre should be further investigated through a DoE approach. With proper planning, this process development strategy can be completed in under 48 hours.

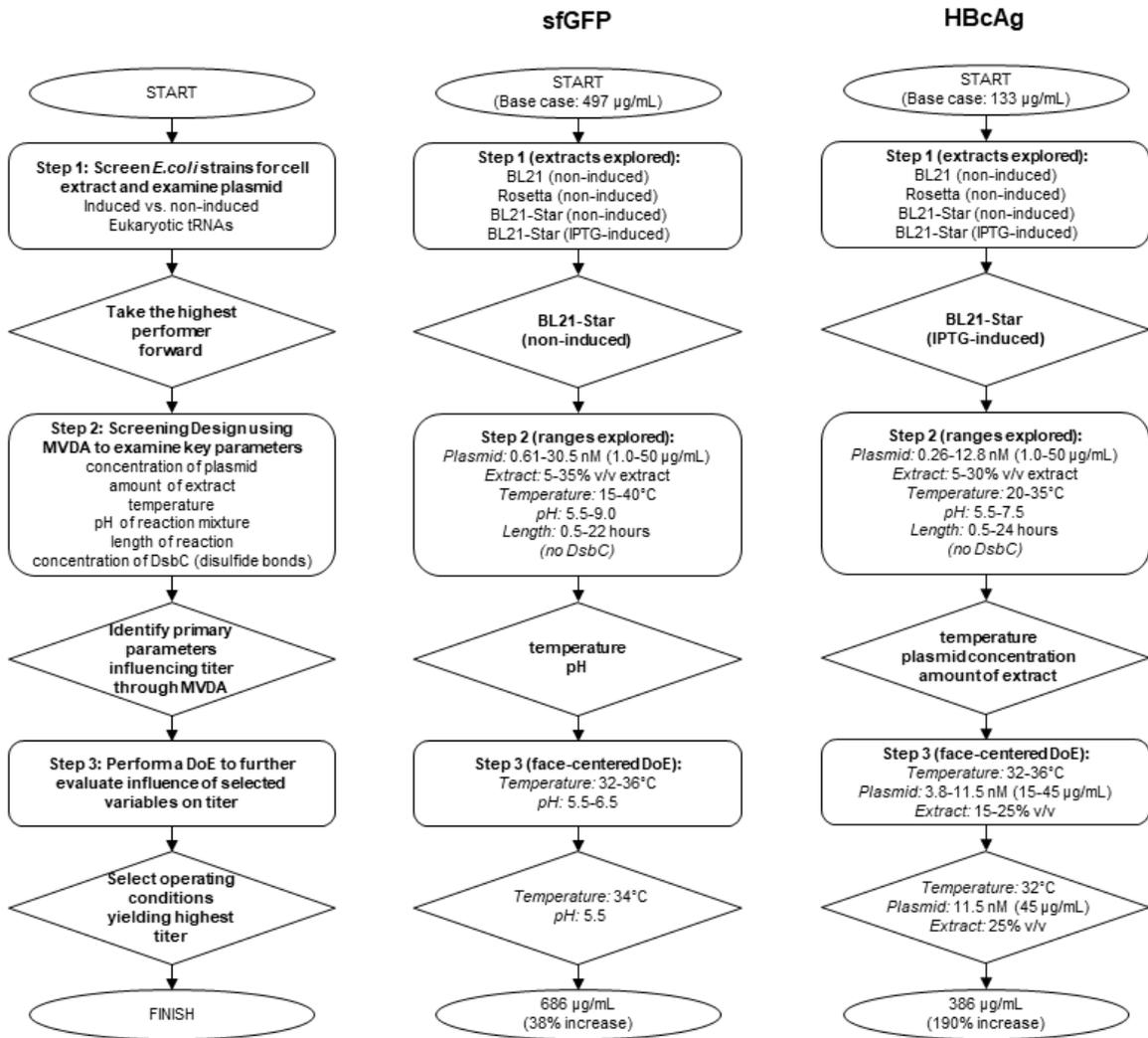


Figure 4.6 Flow Chart of Recommended Process Development Strategy for *E. coli* CFPS reactions

#### 4.6 Conclusion

CFPS allows for recombinant protein production over short reaction times in small reaction volumes; therefore a process development strategy can be deployed that could be completed in under 48 hours. In Chapter 3, the *E. coli* strain chosen for the cell extract was determined to be the most critical parameter. Beyond that, titre can be incrementally increased by manipulating other process parameters. This method can be easily applied to new or difficult to express products in addition to efficiently testing a wide range of reaction conditions. It is recommended to first examine a series of extracts from different *E. coli* strains and to observe the effects of IPTG-induction where possible before conducting further

experiments on other parameters (Figure 4.6). Subsequently, MVDA modelling should be applied to determine the influence of process parameters on production. Then, a DoE study should be used to identify the process conditions that result in the highest titre in the design space. Using this strategy, sfGFP and HBcAg titres were increased by 38% and 190%, respectively, beyond the standard conditions.

This method is recommended over an initial scouting DoE due to the multitude of parameters that can be manipulated in CFPS reactions. Here extract strain, concentrated reaction mixture composition, plasmid selection, plasmid concentration, amount of extract, pH of the concentrated reaction mix, reaction temperature, and length of reaction were examined. Depending on the product, other parameters like chaperone concentration, agitation rate, T7 RNA polymerase concentration, osmolality of the concentrated reaction mix, or protease inhibitor addition could also be critical for improving titre. While an initial DoE study to examine all these parameters could easily be performed given the high-throughput nature of CFPS, many parameters would not be critical to improving titre and the CFPS reactions in which those parameters are manipulated would be a waste of time and resources. By using MVDA to determine the parameters with the highest influence on titre, other parameters may be ruled out and a DoE approach that focuses only on the most critical parameters may be taken.

## 5. Producing AAV2 VLPs in CFPS

*The CFPS system developed in this project has been demonstrated to be capable of producing model products like superfolder GFP (sfGFP) and hepatitis B core antigen (HBcAg) virus-like particles (VLPs). Because one of the goals of this project was producing self-assembling particles with gene therapy applications, adeno-associated virus serotype 2 (AAV2) virus-like particles (VLPs), a product that might be used for gene therapy were chosen for further study. The three capsid proteins, VP1, VP2, and VP3 were expressed in CFPS individually. As an attempt to form capsids, their plasmids were combined in the CFPS reactions; however, this attempt was not successful. The proteins each had a 6x histidine tag which allowed them to be purified individually using Ni-NTA resin, although only a very small amount of material was eluted from the resin. In the future, these purified proteins could be combined in vitro to form capsids. By expressing AAV2 capsid proteins in CFPS, the potential for CFPS to be used for AAV2 capsid formation has been demonstrated which may allow CFPS to be used for gene therapy product and personalised medicines manufacturing.*

### 5.1 Introduction

In the previous chapter, HBcAg VLPs were expressed. In this chapter, expression of a more complex product, adeno-associated virus serotype 2 (AAV2) VLPs was demonstrated. While the HBcAg VLPs are composed of identical monomers, the AAV2 VLPs are composed of three different proteins from the same virus.

AAV2 VLPs include of three capsid proteins: VP1 (87 kDa), VP2 (73 kDa), and VP3 (62 kDa) (Rose et al., 1971). The proteins are all expressed from the same open reading frame (ORF) and share the same C-terminal regions because they are produced from alternatively spliced mRNA (Becerra et al., 1988). In wild

type AAV2, these proteins assemble in a 1:1:10 ratio to form capsids. These capsids have a variety of applications. They have been used to transport siRNA and ssDNA (Shao et al., 2012, King et al., 2001). They have also been modified to display antigen for vaccines, not unlike the tandem-core HBcAg VLPs expressed in the next chapter (Nieto et al., 2012).

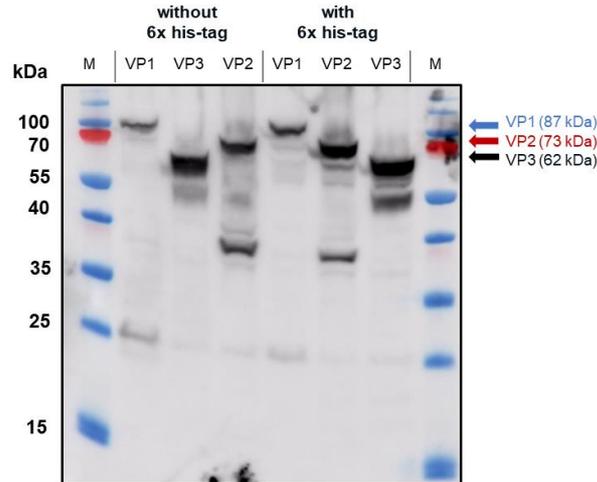
With the demand for viral vectors exceeding the capacity of contract development and manufacturing organisations (CDMOs) and increasing lag times on production, alternative strategies that allow for more rapid generation of gene therapy products are needed (Rininger et al.). AAV2 VLPs produced via CFPS reactions could be a potential solution. AAV2 VLPs have been already been produced in baculovirus cells, HEK293 cells, *Saccharomyces cerevisiae*, and *E. coli* cells (Steinbach et al., 1997, Nieto et al., 2012, Backovic et al., 2012, Le et al., 2019). To the best of my knowledge, this is the first time AAV2 capsid proteins have been expressed in CFPS reactions.

## **5.2 AAV Capsid Protein Production in *E. coli*-based CFPS**

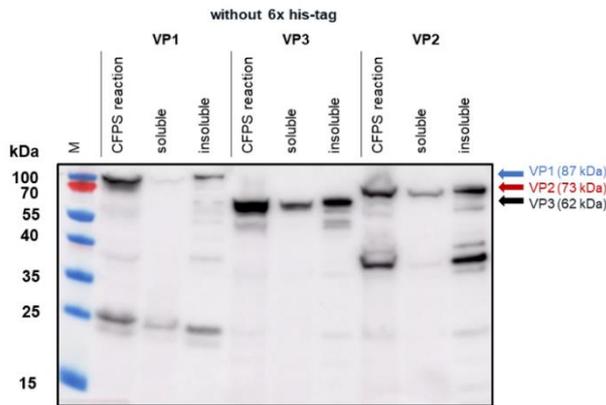
Two constructs for each of the three AAV2 capsid proteins were designed, one with a 6x histidine tag and one without it. The 6x histidine tag was included to enable easier purification of the individual capsid proteins. Similar tags have been used with capsid proteins previously (Wu et al., 2000, Zhang et al., 2002). The six constructs were successfully expressed in 100  $\mu$ L CFPS reactions using the following conditions: 20% IPTG-induced BL21 Star<sup>TM</sup> (DE3) extract, complex concentrated reaction mixture pH 6.0, 10  $\mu$ g/mL of plasmid, 30°C, and 4 hours (Figure 5.1A). Lower molecular weight product related impurities were present as well. However, VP3 proteins previously expressed *in vivo* in *E. coli* have also been accompanied by product related impurities (Le et al., 2019). The reactions were centrifuged, the supernatant was decanted (the soluble fraction) and the pellet was

resuspended in renaturing buffer, 0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7 (the insoluble fraction), and the majority of the material was determined to be insoluble (Figure 5.1B). Again, this is not entirely surprising as VP3 proteins previously expressed *in vivo* in *E. coli* were also largely insoluble (Le et al., 2019). Now that the three proteins have been expressed individually, there are two potential approaches to forming capsids: combining the plasmids for all three plasmids in a single CFPS reaction and individually purifying each protein and assembling the capsids *in vitro*.

A)



B)



C)

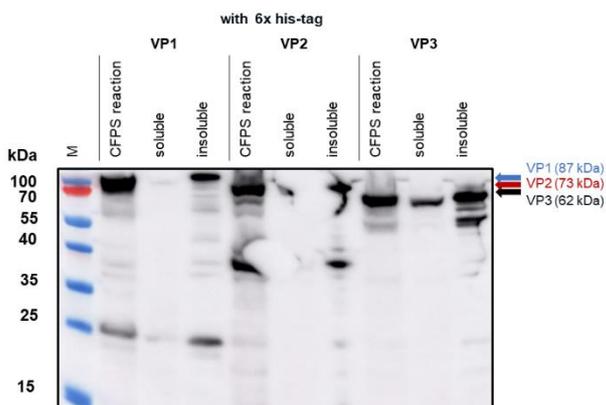


Figure 5.1 CFPS Expression of AAV2 Capsid Proteins

A) AAV2 capsid proteins VP1 (87 kDa), VP2 (73 kDa), and VP3 (62 kDa) with and without a 6x histidine tag were expressed in CFPS reactions under the following conditions: 20% IPTG-induced BL21 Star™ (DE3) extract, complex concentrated reaction mixture pH 6.0, 10 µg/mL of plasmid, 30°C, and 4 hours. B) The reactions containing the capsid proteins without the 6x histidine tag were centrifuged, and the soluble fraction (the supernatant) was separated from the insoluble fraction (the pellet resuspended in renaturing buffer, 0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7). C) The same procedure was performed for reactions containing the capsid proteins with the 6x histidine tag (Note: there was an air bubble during the transfer to the membrane at ~35-40 kDa in the lanes containing the CFPS reaction and the soluble fraction for VP2).

### **5.3 Combining Plasmids for Capsid Protein in CFPS Reactions for VLP**

#### **Formation**

Capsid formation was initially attempted by combining the three plasmids in a single CFPS reaction. Three combinations of the plasmids were tested. The first ratio chosen was 1 VP1: 1 VP2: 10 VP3 because that is the ratio of the three proteins in the native virus. A ratio of 1 VP1: 1 VP2: 1 VP3 was also chosen to determine whether one of the capsid proteins was preferentially expressed in the CFPS system. The third combination did not include any VP1 or VP2, only VP3 because capsids have been formed with VP3 alone previously.

All three capsid proteins were produced in the reactions with plasmid ratios of 1:1:10 and 1:1:1. Although the proteins had been expressed, fully formed capsids were not visible when the samples were imaged under TEM (Figure 5.2). Capsids were not formed in the reactions containing VP3 alone either, although there were small circular protein aggregates visible. These aggregates were much smaller than the expected 22 nm particles. This indicates that the CFPS reaction conditions are not conducive to AAV2 VLP assembly. In the future, changes could be made to the reaction environment that have been shown to help *in vitro* assembly, like the addition of L-arginine or increasing the pH of the reaction (Le et al., 2019).

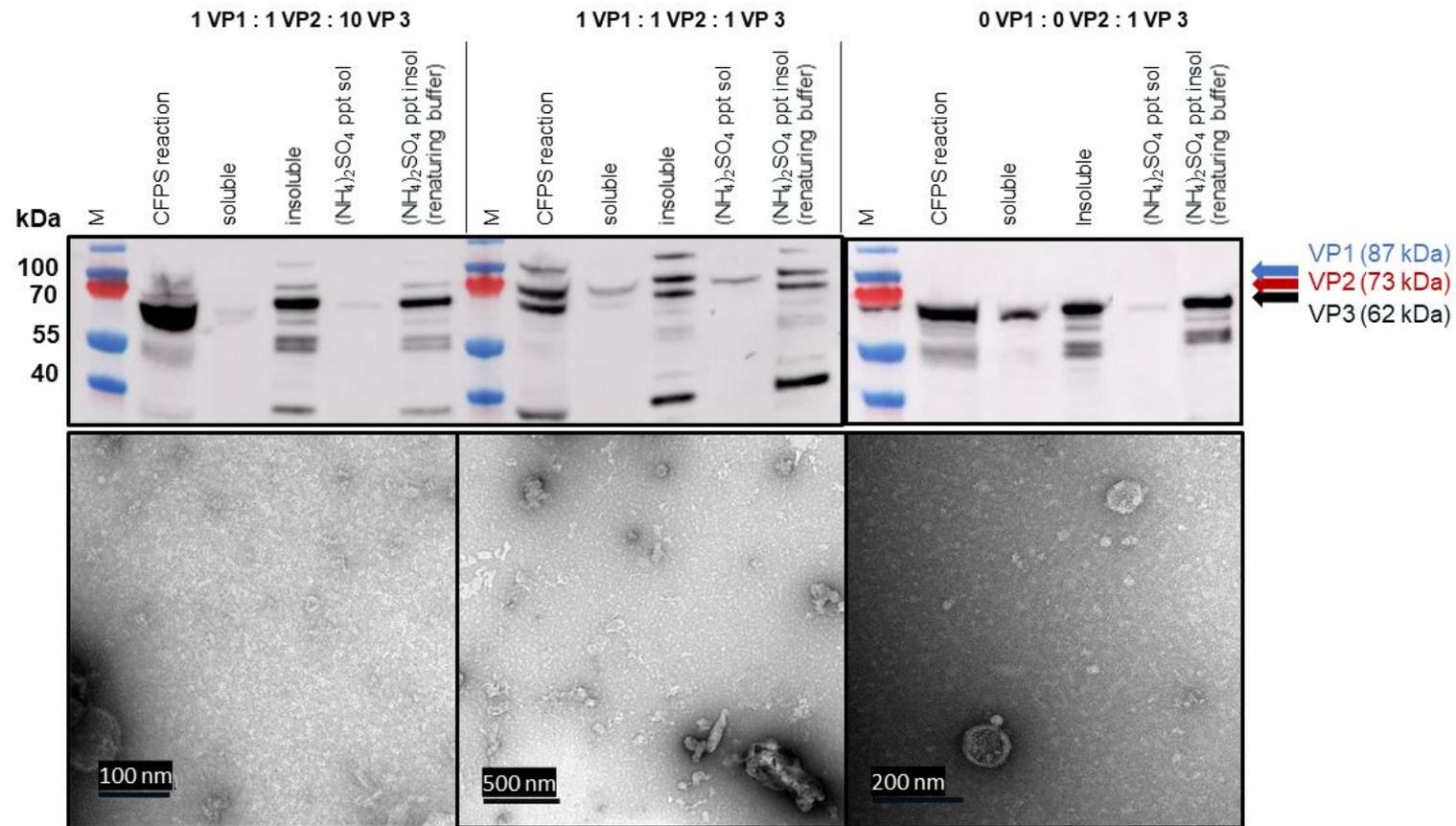


Figure 5.2 Combining Capsid Protein Plasmids in CFPS Reactions for Capsid Assembly

CFPS reactions were performed as previously described with the plasmids for the three AAV2 capsid proteins in a 1:1:10 ratio, the plasmids for the three capsid proteins in a 1:1:1 ratio, or only the plasmid for VP3. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in renaturing buffer (0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7) and saved as the insoluble fraction. The soluble fraction was then precipitated with ammonium sulphate. The supernatant was collected and saved as (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ppt sol. The pellet was resuspended in renaturing buffer and saved as (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ppt insol. (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> ppt insol was imaged under a TEM. No particles were observed in any of the samples, although circular protein aggregates were present in the reactions containing only plasmid for VP3 protein. These TEM images were taken by the author under the supervision of Mark Turmaine at UCL.

#### **5.4 Purifying Capsid Proteins for In Vitro VLP Formation**

Rather than combining the three plasmids in the CFPS reaction to form particles, each capsid protein with a 6x histidine tag was expressed individually and purified them using Ni-NTA resin. While some of the protein was eluted using elution buffer (100 mM Tris, 1 mM EDTA, 500 mM NaCl, 250 mM imidazole, pH 8.7), most of the protein remained on the resin (Figure 5.3 and Table 5.1). This could be because protein was bound strongly to the resin and the elution conditions were not appropriate to remove the protein. This seems unlikely as comparable proteins have been successfully removed using similar elution conditions, but if this is the case, using zinc or cobalt resin with a lower affinity for the 6x histidine tag than a nickel resin may mitigate this issue. Instead, because this protocol involved the centrifugation of the Ni-NTA resin with the bound protein and the protein is largely insoluble, it is more likely that the protein is precipitated with the resin even during elution. Adding substances to the elution buffer to improve solubility, like urea or guanidine-HCl could help improve elution. Using alternative strategies for Ni-NTA purification like gravity flow or fast liquid protein chromatography (FPLC) could avoid precipitation of the proteins.

In future experiments, the proteins purified via Ni-NTA resin could be combined for assembly. Each capsid protein with and without a 6x histidine tag would be expressed individually and isolated via ammonium sulphate precipitation. The capsid proteins would then be solubilised with 5 M guanidine-HCl and 1 mM DTT, combined, and dialysed against PBS with 0.2 M L-arginine pH 9.0 at 4°C as recommended in (Le et al., 2019). The exact conditions for assembly may need to be adjusted; previous work suggests investigating pH, ion strength, and concentration of L-arginine to determine the appropriate conditions (Le et al.,

2019). Properly assembling the VLPs will require testing different ratios of the three capsid proteins and testing different combinations of proteins with and without the 6x histidine tag. Although the tag allows for a more specific purification protocol, another group expressed all three proteins with a 6x histidine tag, but they were not able to assemble into particles (Zhang et al., 2002). This may mean that alternative tags or purification strategies will need to be explored if the proteins cannot be sufficiently isolated via ammonium sulphate precipitation. There are affinity resins specific to AAV on the market that may be of use, although it is unclear how well they bind to each individual capsid proteins. AAV viral vectors have been purified previously using ion exchange chromatography and hydrophobic interaction chromatography; these protocols may be able to be adapted for the capsid proteins (Adams et al., 2020).

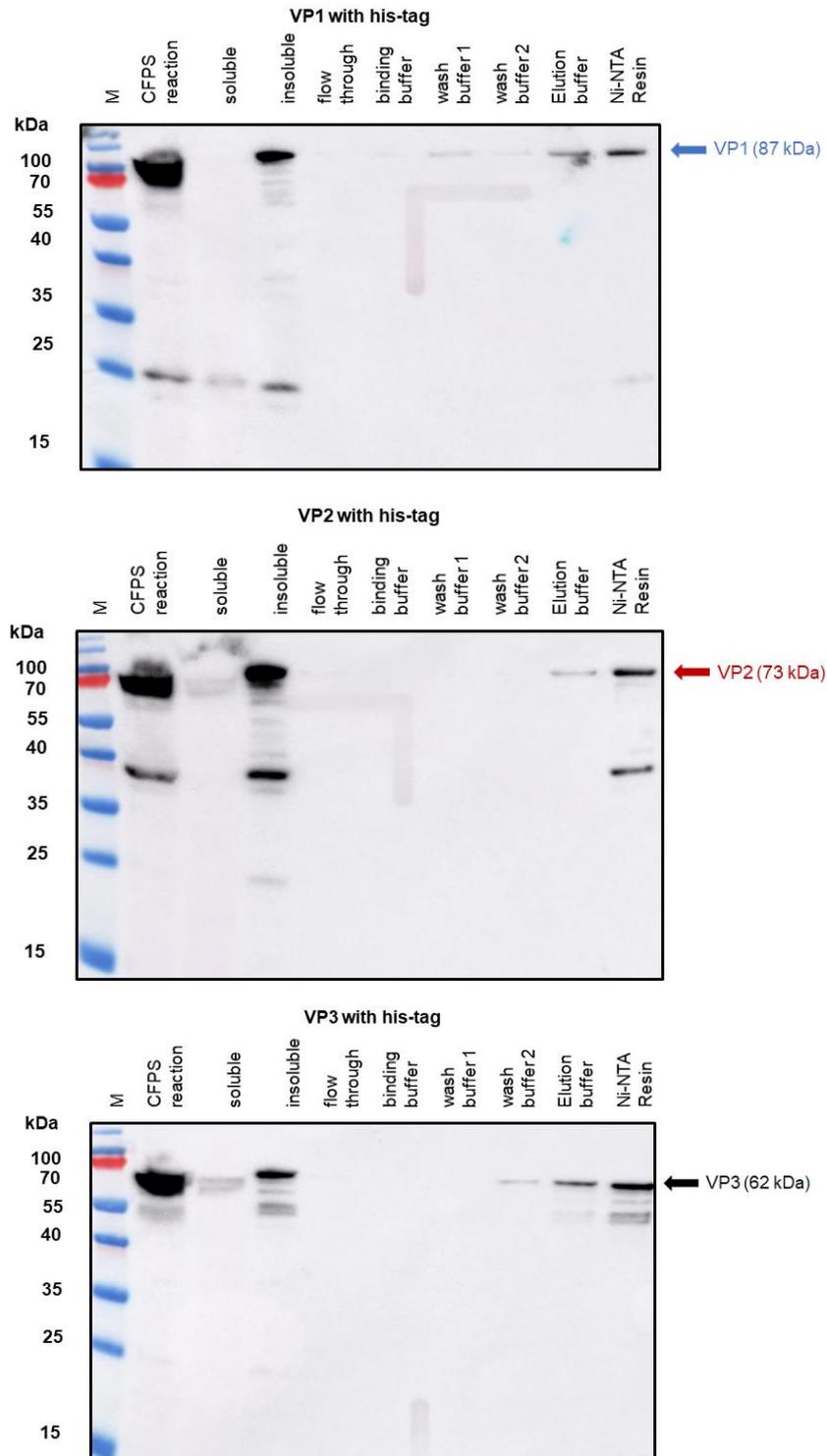


Figure 5.3 Ni-NTA Purification of AAV2 Capsid Proteins

CFPS reactions were performed as previously described with the plasmids for the three AAV2 capsid proteins individually. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in renaturing buffer (0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7) and saved as the insoluble fraction. The insoluble fraction was loaded onto a Ni-NTA resin, washed with binding buffer, two wash buffers, and eluted. A western blot with the [B1] primary antibody was performed on the CFPS reaction, the soluble fraction, the insoluble fraction, the flow through, the wash with binding buffer, two additional washes, the elution, and a sample of Ni-NTA resin.

Table 5.1 Bradford Assay of AAV2 Capsid Proteins Before and After Ni-NTA Purification

	<b>AAV2 VP1 (<math>\mu\text{g/mL}</math>)</b>	<b>AAV2 VP2 (<math>\mu\text{g/mL}</math>)</b>	<b>AAV2 VP3 (<math>\mu\text{g/mL}</math>)</b>
<i>Resuspended pellet from CFPS reaction (before Ni-NTA purification; contains impurities)</i>	478 $\pm$ 49	687 $\pm$ 50	434 $\pm$ 19
<i>Elution fraction following Ni-NTA purification</i>	22 $\pm$ 12	9 $\pm$ 5	15 $\pm$ 14

### 5.5 Conclusion

The work in this chapter established that the three AAV2 capsid proteins, VP1, VP2, and VP3, can be expressed using CFPS. It was also demonstrated that typical CFPS conditions do not allow for *in vitro* assembly of AAV2 capsids. Although adjusting the reaction conditions may result in an environment more conducive to capsid formation, other groups working with *E. coli* have had success expressing individual capsid proteins and then attempting assembly (Le et al., 2019). The individual capsid proteins with 6x histidine tag were expressed, and purification using a Ni-NTA resin was attempted. Some of the material was eluted off the resin, but most of the protein remain precipitated with the resin. By using buffers to improve solubility or by using the Ni-NTA in a column rather than a microcentrifuge tube, these issues may be mitigated, and more protein may be eluted. That material could then be quantified and combined *in vitro* under previously used conditions to form capsids. Once these capsids are formed, they could be loaded with genetic material and potentially used as gene therapy products.

## **6. Characterisation of Product Related Impurities of Tandem-Core Virus-like Particles Produced Using *Escherichia coli*-based CFPS**

*In order to demonstrate that the CFPS system in this project is capable of producing self-assembling vaccine products, it was validated with a series of tandem-core HBcAg VLPs that were designed as universal influenza A vaccine candidates. These vaccine candidates were chosen because they self-assemble, they combine proteins from two different parent viruses (the hepatitis B virus and the influenza A virus), and they have previously been produced in vivo, although they tend to form aggregates. Expression and assembly of the original constructs for two tandem-core vaccine candidates was successful. However, there were several lower molecular weight product related impurities expressed along with the full-length product for both candidates and one candidate was largely insoluble. It was determined that these impurities were not the result of protease cleavages. The plasmid backbone and the gene sequence for the two tandem-core vaccine candidates were then modified to attempt to mitigate these issues. Although the solubility of the vaccine candidates was not able to be improved, the presence of lower molecular weight product related impurities was decreased and the overall titre of the reaction was improved. By expressing these modified tandem-core HBcAg VLP universal influenza A vaccine candidates, the use of CFPS to rapidly iterate upon construct design to improve the quality of self-assembling vaccine candidates has been demonstrated.*

### **6.1 Introduction**

The CFPS manufacturing system in this project has been developed, a process development strategy has been designed, and that strategy has been validated with two model products, superfolder GFP (sfGFP) and hepatitis B core antigen (HBcAg) virus-like particles (VLPs). Here the CFPS manufacturing system

was used to express more complicated self-assembling products for vaccine production. The tandem-core HBcAg VLPs designed as an attempt at a universal influenza A vaccine were chosen.

Most influenza A vaccines need to be redesigned each year because they target regions of the influenza A virus that mutate frequently, neuraminidase and the globular head of haemagglutinin. These vaccines also require several months to produce, so the gene sequences for the viral proteins are predicted ahead of time; poor predictions can result in vaccines that are not well-suited to the emergent strains (Xie et al., 2015). The tandem-core vaccine candidates target more conserved regions of the virus, the matrix 2 ectodomain (M2e) and the haemagglutinin stalk protein (HA2). This should allow for a more stable vaccine that confers protection for several years. The tandem-core VLPs consist of two HBcAg monomers tethered together with a glycine-glycine-serine linker. HBcAg monomers are made of four alpha helices and between the two central alpha helices is a major insertion region (MIR) where other peptides can be inserted. However, peptides that are large or hydrophobic can prevent the VLP from assembling. The linker holding the two monomers together in the tandem-core constructs stabilises the dimer and encourages assembly. In one vaccine candidate, referred to as VLP 1, HA2 was inserted in one MIR and three versions of the M2e protein were inserted in the other. Another simpler vaccine candidate called VLP 3 was made with just the M2e proteins in one MIR. For both products, mice injected with the vaccines generated antibodies against the antigens in the MIRs; vaccination also resulted in an increased survival rate after infection with an influenza A virus, although it did not induce sterilising immunity (Ramirez et al., 2018).

These products were originally manufactured in both *E. coli* and *Pichia pastoris*, but the particles were often aggregated and malformed, especially in *E. coli* (Ramirez et al., 2018, Kazaks et al., 2017). This may be due to steric effects from the influenza inserts. The haemagglutinin stalk protein typically exists as trimer and the matrix 2 ectodomain typically exists as a tetramer. If those protein interactions are occurring, they may be preventing proper particle assembly. Using CFPS provides two potential routes of improving tandem-core VLP expression. First, by expanding the reaction conditions that are used, a set of conditions not achievable *in vivo* may be found that allows for the generation of more consistent, properly assembled product. Second, rapid iteration on the design of the tandem-core VLPs themselves can be undertaken to quickly determine which modifications result in improved products.

## **6.2 Producing Tandem-Core VLPs in an *E. coli*-based CFPS System**

The CFPS manufacturing system described in previous chapters was used to express four tandem-core products. The following conditions were used for all CFPS reactions in this chapter unless otherwise stated: 20% IPTG-induced BL21 Star™ (DE3) extract, complex concentrated reaction mixture pH 6.0, 10 µg/mL of plasmid, 30°C, and 4 hours. The products expressed were K1K1, in which there are no antigens in the HBcAg MIRs – only lysine residues, VLP 1, in which there is an HA2 antigen in the first MIR and three M2e variants in the second MIR, VLP3, in which there are three M2e variants in the first MIR, and tcHBcAg-GFP, in which wild type GFP is present in the second MIR. All four tandem-core products were successfully expressed in the system and detected on a western blot (Figure 6.1).

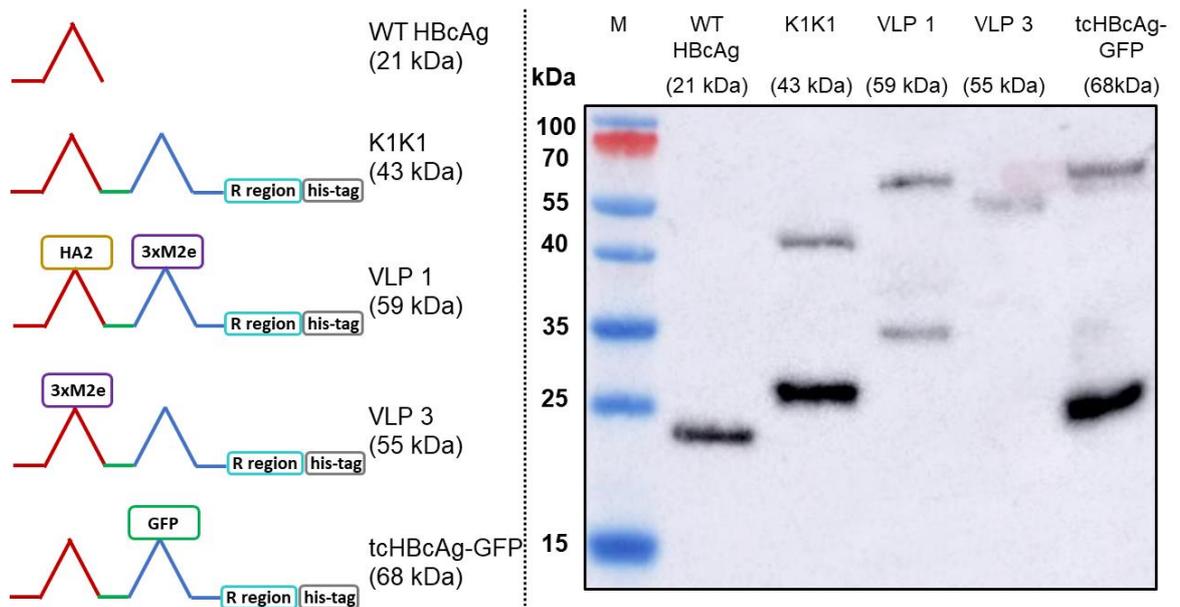


Figure 6.1 CFPS Expression of Tandem-Core HBcAg VLP Constructs

The wild type hepatitis B core antigen (WT HBcAg) along with four tandem-core constructs were successfully expressed in the CFPS system. The tandem-core constructs were K1K1, which has lysine residues in both of its MIRs, VLP 1, which has haemagglutinin stalk protein HA2 in its first MIR and three M2e proteins in its second MIR, VLP 3, which has three M2e proteins in its first MIR, and a tandem-core constructs with GFP in its second MIR (tcHBcAg-GFP). A western blot with the [14E11] primary antibody was performed on all five proteins. The top band in each lane is the full-length product.

The tandem-core HBcAg VLPs with the influenza antigens (VLP 3 and VLP 1) are of greatest interest to this project and were taken forward. Unfortunately, both products were largely insoluble when produced under typical conditions in the CFPS system. This is not entirely unexpected as tandem-core VLPs produced in *E. coli* have been shown to precipitate (Holmes et al., 2015). Decreasing the temperature of the *in vivo* cultures can prevent the formation of insoluble inclusion bodies (Schumann and Ferreira, 2004). The temperature of the reactions for both products was decreased. When VLP 3 was produced in reactions at 18°C, there was less insoluble product detected (Figure 6.2). This was not the case for VLP 1. When VLP 1 was produced at temperatures below 21°C, the material was still insoluble and very little full-length product was present in the reaction (Figure 6.3).

All subsequent reactions for VLP 3 and VLP 3 derived products were performed at 18°C and reactions for VLP 1 and VLP 1 derived products were performed at 30°C.

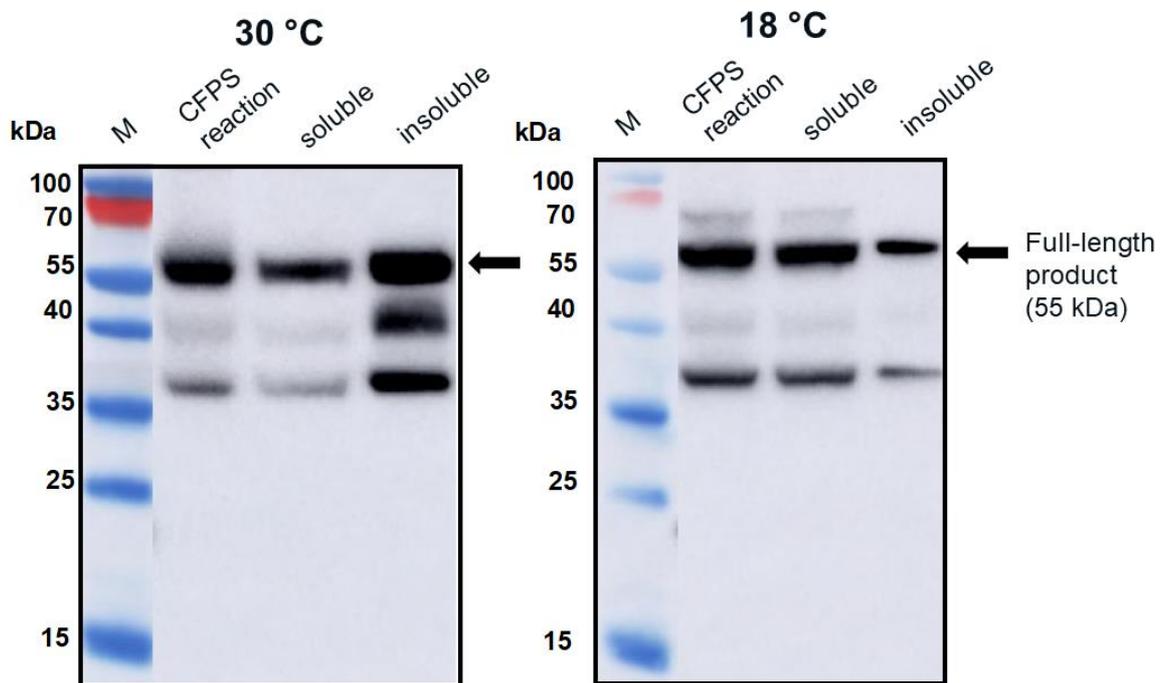


Figure 6.2 CFPS Reactions to Produce VLP 3 at 30°C and 18°C

CFPS reactions to produce VLP 3 were performed at both 30°C and 18°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. A western blot with the [10E11] primary antibody was performed on the CFPS reaction, the soluble fraction, and the insoluble fraction. At 30°C, the material in the reaction is mostly insoluble. At 18°C, more VLP 3 can be found in the soluble fraction.

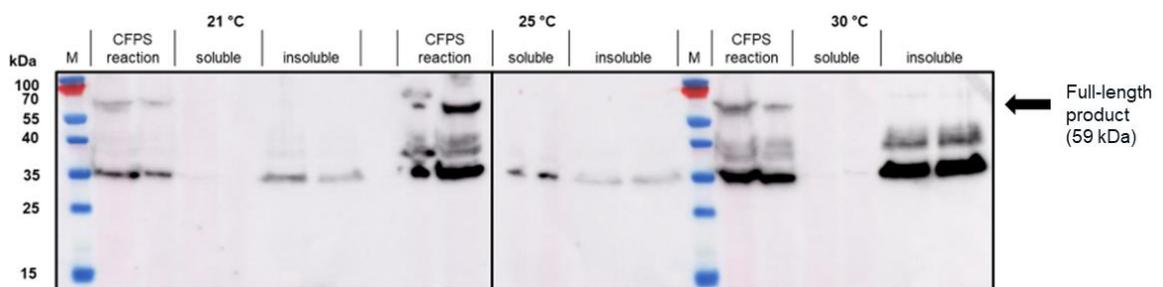


Figure 6.3 CFPS Reactions to Produce VLP 1 at 21°C, 25°C and 30°C

Duplicate CFPS reactions to produce VLP 1 were performed at 21°C, 25°C, and 30°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. A western blot with the [10E11] primary antibody was performed on the CFPS reaction, the soluble fraction, and the insoluble fraction. At all three temperatures, the VLP 1 material is largely insoluble. At 21°C, there is barely any full-length material in the reaction.

It was hypothesised that VLP 3 and VLP 1 might be insoluble because they were aggregating with host cell proteins present in the cell extract. If the host cell proteins could be removed, it is possible that the tandem-core VLPs may be soluble. Because both the VLPs have a C-terminal 6x histidine tag, purification of the particles using Ni-NTA resin was attempted. The particles did not bind to the resin at all (data not shown). The 6x histidine tag has been shown to stabilise HBcAg VLPs and related products because the tags join together inside of the VLP and prevent the monomer, or in this case tandem-core dimers, from separating (Schumacher et al., 2018). The linkers used in the tandem-core products also increase stability. This means that monomers or tandem-core dimers that have already come together to form particles or aggregates will not display the 6x histidine tag, as it is inside of the particle. Additionally, these particles and aggregates are very difficult to separate. Another purification using Ni-NTA resin and buffer containing 8 M urea to separate the particles and aggregates was attempted. Again, the material did not bind to the resin which suggests that even denaturants like 8 M urea cannot separate the tandem-core dimers once they bind together. Other research groups have demonstrated that tandem-core HBcAg VLPs are more resistant to denaturants than wild type HBcAg VLPs (Holmes et al., 2015). It might be possible to separate the dimers by fusing another protein to tandem-core dimers that can also be used as an affinity tag, like maltose-binding protein (MBP) or glutathione S-transferase (GST) (García-Fruitós, 2015). Other proteins like thioredoxin (TrxA), DsbC, small ubiquitin-like modifier protein (SUMO), and N-utilisation substance A (NusA) can increase solubility and can be used in conjunction with smaller affinity tags, like the 6x histidine tag (García-Fruitós, 2015). These tags, which are large enough to prevent particles from self-

assembling, will need to be removed following purification so that the tandem-core VLPs can properly assemble.

### **6.3 Observing Tandem-Core Product Related Impurities**

VLP 3 and VLP 1 were expressed, precipitated using ammonium sulphate, and imaged under the TEM (Figure 6.4 and Figure 6.5). Although individual symmetrical particles were observed for both products, aggregated particles and large proteins were also present. It is likely that many of the large proteins are residual host cell proteins from the cell extract and may be able to be removed with a more sophisticated purification strategy. However, there may be some VLP material in the aggregate. In future studies, immunogold staining could be used to determine whether a substantial amount of that aggregated material included tandem-core VLPs (Stephen et al., 2018). Because the VLPs are insoluble, especially VLP 1, it is not entirely surprising that aggregate particles are present. The aggregates may also be the result of the lower molecular weight product related impurities observed on the western blot analysis. In addition, while the VLP 3 particles are the expected size and shape, the particles formed by the VLP 1 construct were larger than anticipated, ~60 nm in diameter. Based on previous work expressing this construct in *Pichia pastoris*, particles that are ~40 nm in diameter are expected (Kazaks et al., 2017). Perhaps a larger conformation is preferable in CFPS or proteins from the CFPS reaction are inside of the particles or inhibiting the formation of smaller particles. It is also possible that due to the fact that VLP 1 has two different influenza inserts (HA2 and M2e) those proteins are interacting with other inserts on other VLP 1 dimers and the resulting steric effects are preventing formation of particles with the proper diameter.

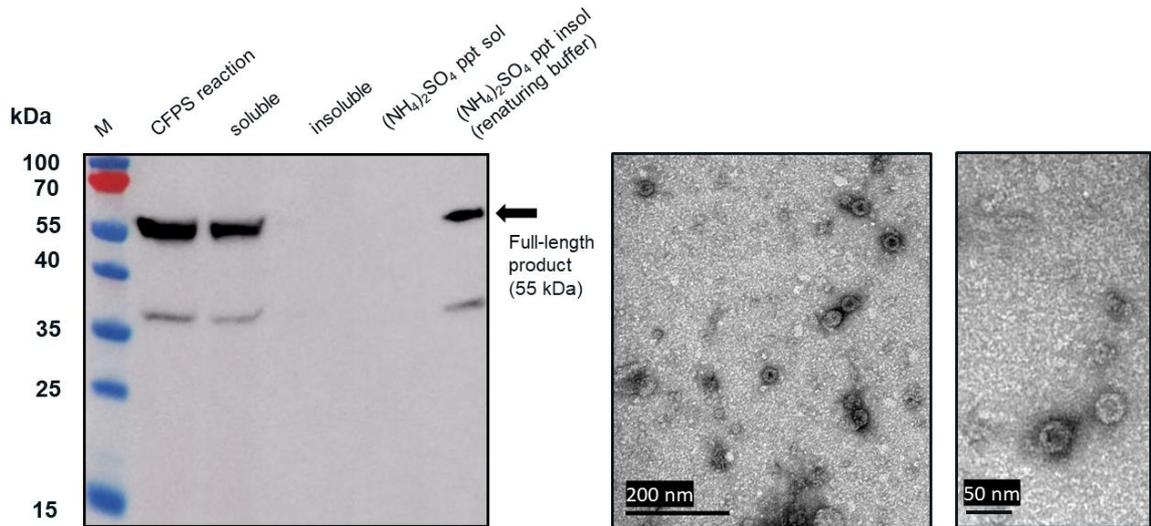


Figure 6.4 Ammonium Sulphate Precipitation and TEM images of VLP 3

CFPS reactions to produce VLP 3 were performed at 18°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. The soluble fraction was then precipitated with ammonium sulphate. The supernatant was collected and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt sol. The pellet was resuspended in renaturing buffer (0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7) and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt insol. A western blot with the [14E11] primary antibody was performed on the CFPS reaction, the soluble fraction, the insoluble fraction,  $(\text{NH}_4)_2\text{SO}_4$  ppt sol, and  $(\text{NH}_4)_2\text{SO}_4$  ppt insol.  $(\text{NH}_4)_2\text{SO}_4$  ppt insol was imaged under a TEM. Symmetrical particles ~40nm in size are visible. These TEM images were taken by the author under the supervision of Mark Turmaine at UCL.

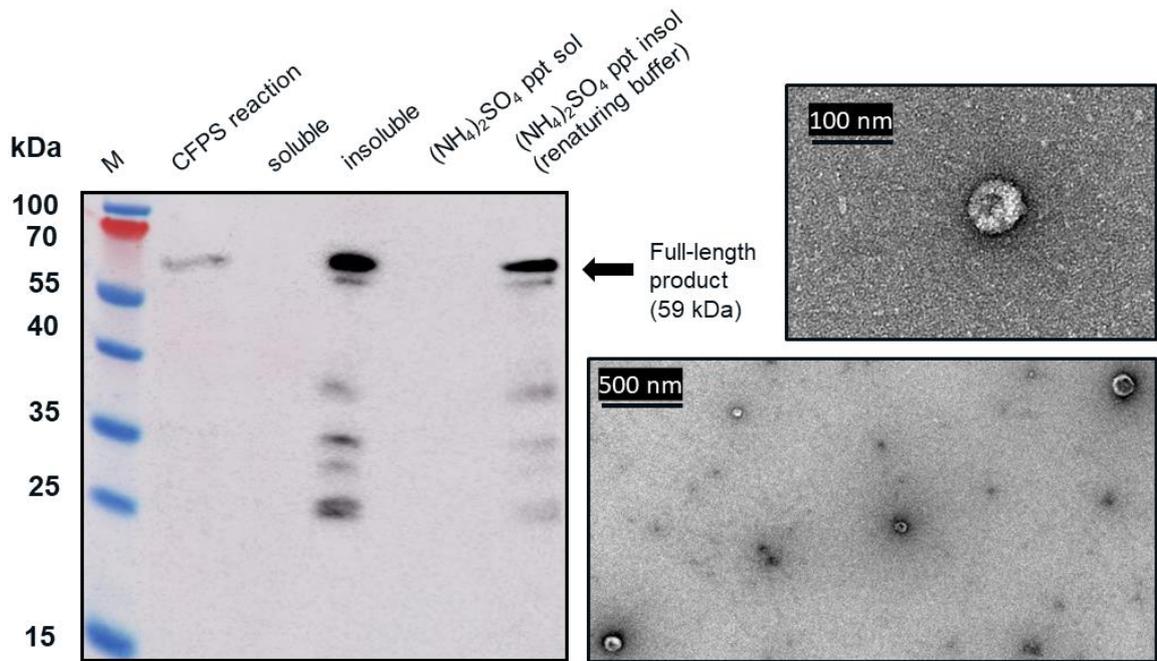


Figure 6.5 Ammonium Sulphate Precipitation and TEM images of VLP 1

CFPS reactions to produce VLP 1 were performed at 30°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. The insoluble fraction was then precipitated with ammonium sulphate. The supernatant was collected and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt sol. The pellet was resuspended in renaturing buffer (0.1 M Tris, 1 mM EDTA, 0.15 M NaCl, pH 8.7) and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt insol. A western blot with the [14E11] primary antibody was performed on the CFPS reaction, the soluble fraction, the insoluble fraction,  $(\text{NH}_4)_2\text{SO}_4$  ppt sol, and  $(\text{NH}_4)_2\text{SO}_4$  ppt insol.  $(\text{NH}_4)_2\text{SO}_4$  ppt insol was imaged under a TEM. Symmetrical particles ~60nm in size are visible. These TEM images were taken by the author under the supervision of Mark Turmaine at UCL.

Product related impurities of ~37 kDa for VLP 3 and ~37 kDa, ~35 kDa, and ~25 kDa for VLP 1 were observed. Similar impurities had been present in the *in vivo* expression of these products (Blaha, 2019). These impurities were presumed to be protease cleavages, and, in the *in vivo* cultivations, protease inhibitor was added to prevent degradation of the full-length product (Blaha, 2019). However, when the cOmplete™ Protease Inhibitor Cocktail was added to the CFPS system, no significant changes in the presence of the product related impurities were seen (Figure 6.6). This led to the conclusion that the product related impurities were not caused by proteases in this CFPS system.

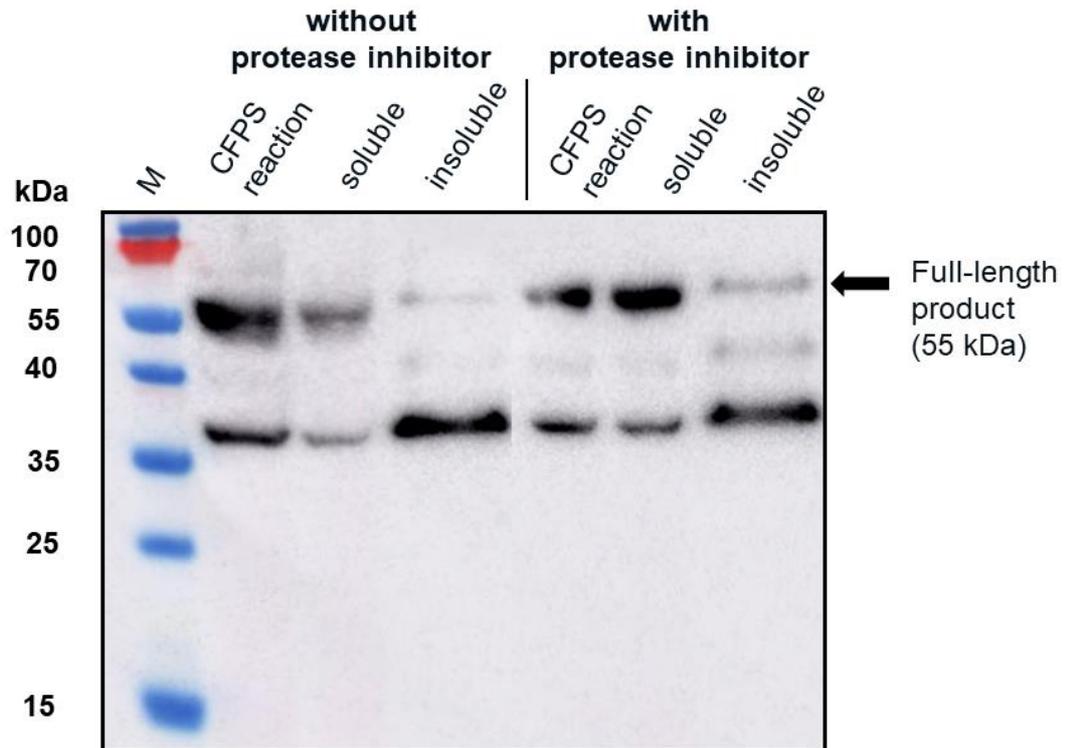


Figure 6.6 Protease Inhibitor Effects on Production of VLP 3

CFPS reactions to produce VLP 3 were performed at 18°C with and without cComplete™ Protease Inhibitor Cocktail. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. A western blot with the [10E11] primary antibody was performed on the CFPS reaction, the soluble fraction, and the insoluble fraction. There is no dramatic difference in solubility or impurity profile when the protease inhibitor is used.

To better understand the nature of these product related impurities, CFPS reactions producing VLP 1 were tested with four different antibodies (Figure 6.7). Each antibody binds to a different region on VLP 1. Anti-Hepatitis B Virus Core Antigen antibody [10E11] (ab8639), which corresponds to amino acids 1-10 on the HBcAg monomer, Anti-Hepatitis B Virus Core Antigen antibody [14E11] (ab8638), which corresponds to amino acids 135-141 on the HBcAg monomer, Anti-Influenza A Virus M2 Protein antibody [14C2] (ab5416), which corresponds to the N-terminal of the Influenza A Virus M2 Protein, and Anti-6X His tag® antibody [HIS.H8] (ab18184), which corresponds to any 6x histidine tag were used. By examining the effects of the four different antibodies, it was confirmed that protease activity was not the cause of these impurities. If the impurities were a result of protease

cleavage, bands of similar molecular weight would be expected no matter which antibody was used. Instead, the full-length product is detected with all antibodies and product related impurities are detected with all antibodies except for [HIS.H8], the antibody corresponding to the 6x histidine tag. In fact, only one band is present with the [HIS.H8] antibody. If the product were cleaved, a lower molecular weight species with the 6x histidine tag would be expected. For example, if the band at ~35 kDa detected with the [10E11] antibody, which indicates that lower molecular weight species has a region corresponding to the beginning of the HBcAg protein, was the result of a cleavage product, then a band at ~25 kDa with the [HIS.H8] antibody would also be expected because that would be the other half of the product. Because there are no other bands on the western blot with the [HIS.H8] antibody, this indicates that the impurities are shorter versions of the full-length product, rather than cleavage products.

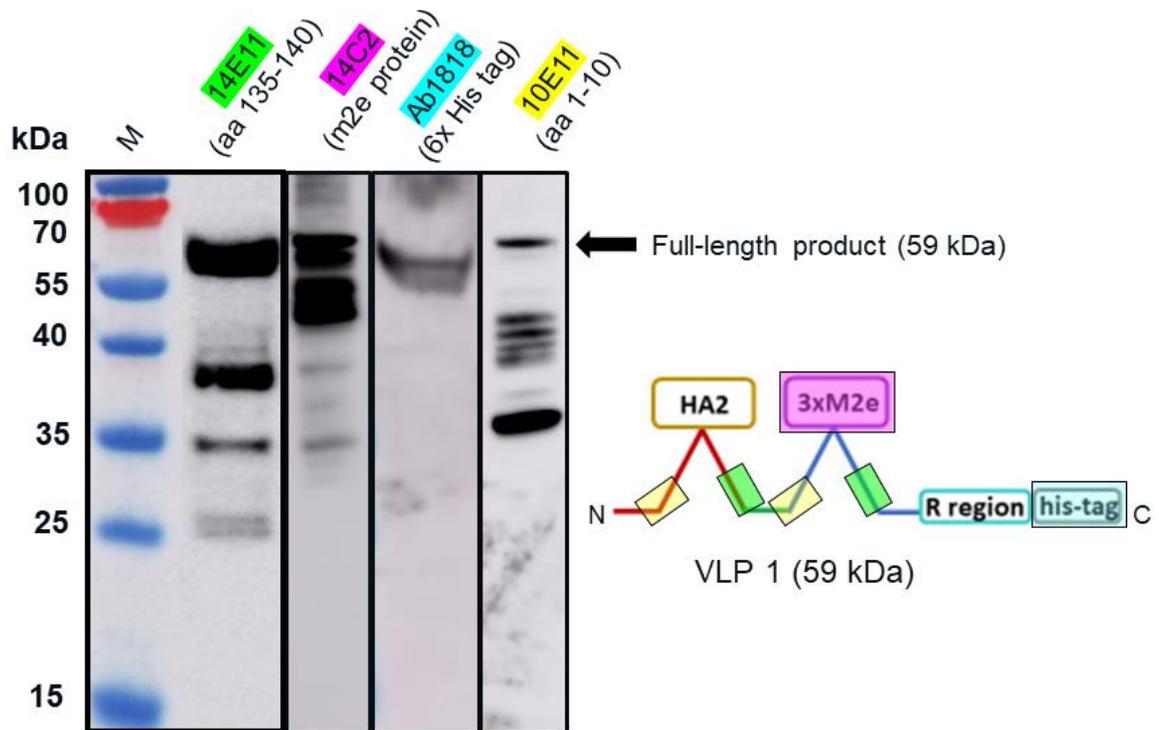


Figure 6.7 Western Blot Analysis of CFPS reaction to product VLP 1 Using Four Different Primary Antibodies

CFPS reactions to produce VLP 1 were performed at 30°C. A western blot was performed on the reactions using Anti-Hepatitis B Virus Core Antigen antibody [14E11] (ab8638), which corresponds to amino acids 135-1414 on the HBcAg monomer, Anti-Influenza A Virus M2 Protein antibody [14C2] (ab5416), which corresponds to the N-terminal of the Influenza A Virus M2 Protein, Anti-6X His tag® antibody [HIS.H8] (ab18184), which corresponds to any 6x histidine tag, and Anti-Hepatitis B Virus Core Antigen antibody [10E11] (ab8639), which corresponds to amino acids 1-10 on the HBcAg monomer. While full-length VLP 1 is detected using all the antibodies, a different impurity profile is observed with each antibody.

It was also concerning that the bands on the western blot that correspond to the product related impurities often appeared much darker than those for the full-length product, implying that there was a greater presence of impurity than product in the reaction. An SDS-PAGE analysis of the material was performed and it was determined that the opposite was true (Figure 6.8). Bands that corresponded to the full-length product on the western blot lined up with similarly dark bands on the Coomassie-stained SDS-PAGE gel. The product related impurities, particularly the band at ~35 kDa visible on the western blot using the [10E11] antibody that appeared much darker than the full-length band, did not show up as well on the Coomassie-stained SDS-PAGE gel. There are two possible explanations: the

impurities are an artifact of the antibodies used for the western blot or the antibodies used for western blot have a stronger preference for the lower molecular weight species. Because different impurity profiles have been observed with different antibodies, the latter is more likely than the former. The antibody may favour the lower molecular weight species because the site it binds to is more accessible in those species than it is in the full-length product, but because the structure of the proteins following denaturation with SDS is not known, that claim is not absolutely certain.

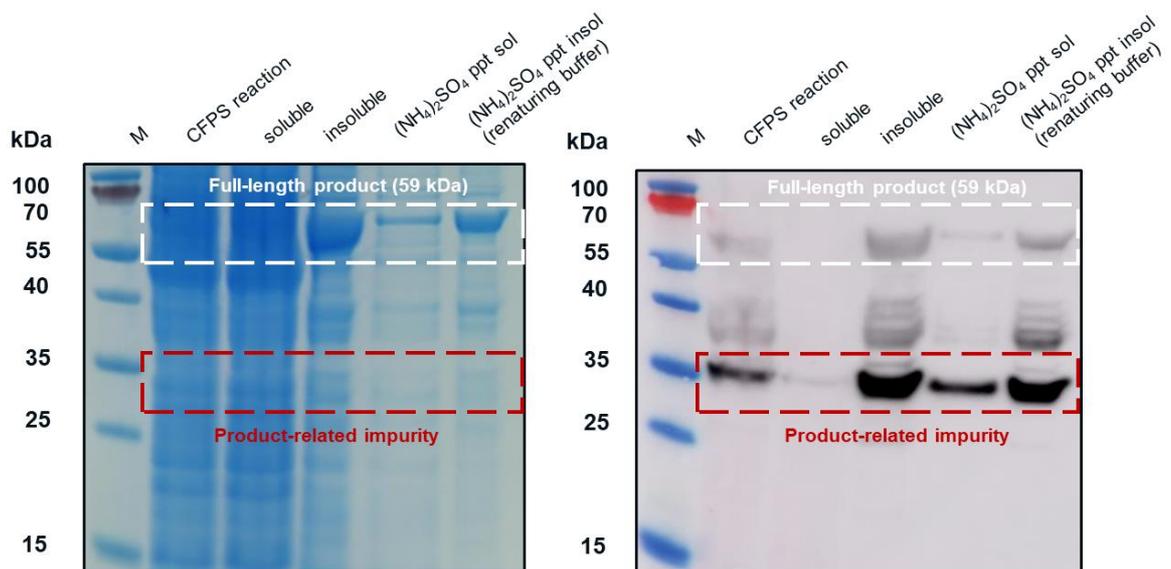
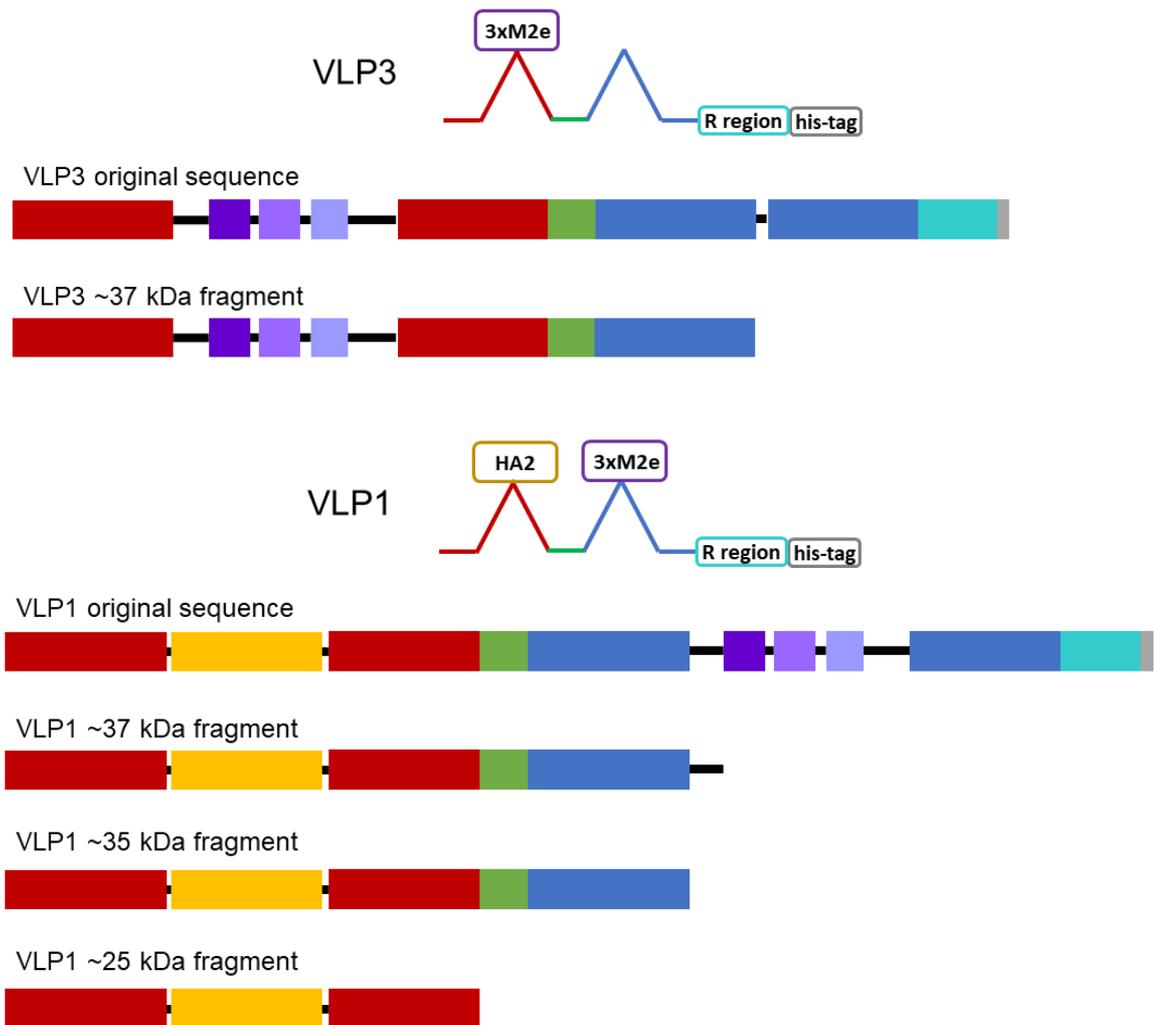


Figure 6.8 SDS-PAGE and Western Blot Analysis of Ammonium Sulphate Precipitation of VLP 1

CFPS reactions to produce VLP 1 were performed at 30°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in 1 M phosphate buffer pH 7.0 and saved as the insoluble fraction. The insoluble fraction was then precipitated with ammonium sulphate. The supernatant was collected and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt sol. The pellet was resuspended in renaturing buffer and saved as  $(\text{NH}_4)_2\text{SO}_4$  ppt insol. An SDS-PAGE analysis and a western blot with the [10E11] primary antibody were performed on the CFPS reaction, the soluble fraction, the insoluble fraction,  $(\text{NH}_4)_2\text{SO}_4$  ppt sol, and  $(\text{NH}_4)_2\text{SO}_4$  ppt insol. Dark bands are visible at ~59 kDa where VLP 1 is expected on both the SDS-PAGE and the western blot (shown in the white box). Dark bands are visible on the western blot at ~35 kDa but are not present on the SDS-PAGE (shown in the red box).

Now that it is known that the impurities are present at lower concentration than they first seemed and that they are likely not caused by protease cleavage, it is likely that the lower molecular weight product related impurities are a result of

partial transcription of the plasmid DNA or partial translation or degradation of the mRNA. The size of lower molecular weight product related impurities (~37 kDa for VLP 3 and ~37 kDa, ~35 kDa, and ~25 kDa for VLP 1) were estimated and mapped out operating under the assumption that all impurities begin at the first start codon after the ribosome binding sequence in the plasmid (Figure 6.9). It appears that the MIR of the second HBcAg dimer is a particularly troublesome region and that the ~37 kDa fragment for VLP 3 and the ~37 kDa and ~35 kDa fragments for VLP 1 are the result of abortive transcription or translation in that region. The ~25 kDa fragment for VLP 1 is most likely the result of issues with the peptide linker holding the two HBcAg monomers together, although it is a bit surprising that a similar fragment for VLP 3 is not also seen. Abortive transcripts are not uncommon in systems that utilise the T7 RNA polymerase to transcribe products that are expressed under the T7 promoter, although these mRNA sequences are typically only a few nucleotides long (Gong and Martin, 2006). It is also possible that one of the essential components for translation, like tRNAs or ribosomes, is limited in this CFPS system. However, these lower molecular weight product related impurities are present *in vivo* in *P. pastoris* as well as *E. coli*, which may indicate that the impurities are a result of the design of the plasmid containing the gene for the product or the product itself (Blaha, 2019). With that in mind, the plasmids and the gene sequences were redesigned for both VLP 3 and VLP 1.



**Figure 6.9 Lower Molecular Weight Product Related Impurities for Tandem-Core HBcAg VLPs**

The first HBcAg monomer is represented with red boxes, the second HBcAg monomer is represented with blue boxes, the haemagglutinin stalk protein is represented as a yellow box, the three M2e proteins are represented as purple boxes, the arginine-rich region is represented as an aqua box, the 6x histidine tag is represented as a grey box, the linker holding the two monomers together is represented as a green box, and all other linkers are represented as black lines. Based on the bands visible on the western blot analyses, VLP 3 has one lower molecular weight product related impurity ~37 kDa in size. This fragment could be the result of abortive transcription or translation in the MIR of the second HBcAg monomer. VLP 1 has three lower molecular weight product related impurities ~37 kDa, ~35 kDa, and ~25 kDa in size. The ~37 kDa fragment likely extends only to the linker holding the M2e proteins in the MIR of the second HBcAg monomer, the ~35 kDa fragment likely extends only to the beginning of the MIR of the second HBcAg monomer, and the ~25 kDa fragment contains only the first HBcAg monomer with the HA2 protein in its MIR.

## **6.4 Examining the Impact of Plasmid Backbone and Gene Sequence**

### **Changes**

VLP 3 and VLP 1 are both hetero-tandem core constructs. That means that the first HBcAg molecule in the construct does not have the nucleic acid binding arginine-rich region that is present in wild type HBcAg, but the second HBcAg molecule in the construct does (see Figure 1.6C for a diagram). Other groups have intentionally removed the arginine rich region to improve stability and prevent lower molecular weight impurities (Walker et al., 2008). Two constructs were designed, one for VLP 3 and one for VLP 1, where the arginine-rich region for each product was removed, but the plasmid and gene sequence were otherwise unchanged. These were called “VLP 3 pET28b no arg” and “VLP 1 pET28b no arg”.

Plasmids were also designed where the gene for each of the tandem-core VLPs (with the arginine-rich region as originally designed) was in the pJL1 plasmid backbone which was used previously to express sfGFP in Chapter 3 and Chapter 4. This backbone is much smaller (1766 bp) than the pET28b backbone (5209 bp) originally used for the tandem-core HBcAg VLPs. The pJL1 plasmid backbone has been optimised for use in CFPS reactions, which should allow for better expression. These constructs were called “VLP 3 pJL1” and “VLP 1 pJL1”.

Finally, plasmids were designed for both VLP 3 and VLP 1 in which the gene sequence was placed in the pJL1 backbone, the arginine rich region was omitted from the sequence, and the peptide sequences linking the HBcAg fragments to the influenza A antigens were made more uniform in length. Based on the length of the truncated fragments observed on the western blots, it seemed likely that the peptide linkers in the second MIR could be a potential point for stalled transcription or translation. These constructs were called “VLP 3 pJL1 + linkers” and “VLP 1 pJL1 + linkers”.

Annotated versions of the protein sequences and diagrams of the plasmid backbones can be found in Appendix A. The six modified constructs were expressed in the CFPS system under the following conditions: 20% IPTG-induced BL21 Star<sup>TM</sup> (DE3) extract, complex concentrated reaction mixture pH 6.0, 5 nM of plasmid, 18°C (VLP 3)/ 30°C (VLP 1), and 4 hours. They were then precipitated with ammonium sulphate (the soluble fraction was precipitated for VLP 3 variants and the insoluble fraction was precipitated for VLP 1 variants) and imaged using TEM (Figure 6.10 and 6.11). They were also analysed via SDS-PAGE, western blot, dot blot, and Bradford assay to determine titre (Figure 6.12 and 6.13). A densitometry reading of an SDS-PAGE analysis was performed for each sample (Figure 6.14A). The bands on the western blot were aligned with the bands on the SDS-PAGE analysis and the density of that band was used to calculate the estimated product titre from the Bradford assay results; this process is explained in Section 2.5.6 (Figure 6.14B). The samples were also analysed by dot blot and compared to a standard curve of recombinant HBcAg samples of known concentration (Figure 6.14C). It should be noted that the titres for the dot blot analysis are slightly higher than those estimated using densitometry because lower molecular weight product related impurities also contribute to the signal on the dot blot analysis.

Unfortunately, the new variants of VLP 1 were still insoluble, much like the original VLP 1 construct. However, modifying the constructs did have a few advantages. Changing the plasmid backbone had no noticeable effect on the product related impurities, but it did result in an increase in titre. Titres for VLP 1 pJL1 were at least 1.8 times greater than VLP 1 and titres for VLP 3 pJL1 were at least 1.4 times greater than VLP 3. As seen previously, VLP 1 and VLP 1 pJL1 form particles that are larger than anticipated. In the resuspended precipitate from

reactions with VLP 3 and VLP 3 pJL1 shown in Figure 6.10 A and B, particles at the expected size of ~40 nm are observed.

In the set of reactions shown in Figure 6.10, there do not appear to be any lower molecular weight product related impurities for VLP 3 pET28b no arg or VLP 3, but the bands are very faint indicating that expression was low in general. Lower molecular weight product related impurities were still present in reactions expressing VLP 1 pET28b no arg, but on the western blot analysis, those bands appear fainter compared to the bands for the full-length product. Given that the antibodies used for western blots analyses of these products sometimes show a preference for products of a certain length, there may not be fewer product related impurities in the VLP 1 pET28b no arg samples, but it is a possibility. Removing the arginine-rich region also increases the titre; titres for VLP 1 pET28b no arg were at least 2.5 times higher than titres for VLP 1 and titres for VLP 3 pET28b no arg were at least 1.3 times higher than titres for VLP 3. However, titres for VLP 3 pET28b no arg were not able to be obtained using dot blot analysis. The primary antibody used for the western blot, [14E11], corresponds to amino acids 135-141 on the HBcAg monomer. The arginine rich region begins at amino acid 149 in the HBcAg monomer. It is possible that by removing the arginine-rich region, the folding of the protein has changed such that the amino acids corresponding to the [14E11] antibody are less accessible. This would also explain the rather faint band seen on the western blot analysis of VLP 3 pET28b no arg. One might expect to see that same issue with VLP 1 pET28b since the arginine-rich region is removed in that construct as well. Then again, VLP 1 pET28b no arg has inserts in both of its MIRs while VLP 3 pET28b only has inserts in the first MIR and that distinction may account for the difference in detection with each construct. Also, the particles formed from the constructs without the arginine-rich region appeared ill-formed and

aggregated. This was also the case with the tandem-core HBcAg constructs without the arginine-rich region previously expressed in *E. coli* (Peyret et al., 2015).

The VLP 1 pJL1 + linkers and VLP 3 pJL1 + linkers constructs resulted in samples with little to no detectable lower molecular weight product related impurities. These constructs also resulted in dramatically improved titres; the VLP 1 pJL1 + linkers construct produced at least 4.0 times as much product as the original construct and the VLP 3 pJL1 + linkers construct produced at least 3.4 times as much. The particles from the VLP 1 pJL1 + linkers and VLP 3 pJL1 + linkers constructs were very inconsistent; some particles were symmetrical, and some were aggregated or only partially formed. This is likely due to the removal of the arginine-rich region. Based on the results from the three modified constructs above, the next iteration on the VLP 1 and VLP 3 constructs would be the VLP 1 pJL1 + linkers and VLP 3 pJL1 + linkers constructs with the arginine-rich region as originally designed. Theoretically, these constructs should have fewer assembly issues and they may also have fewer lower molecular weight product related impurities and increased titres.

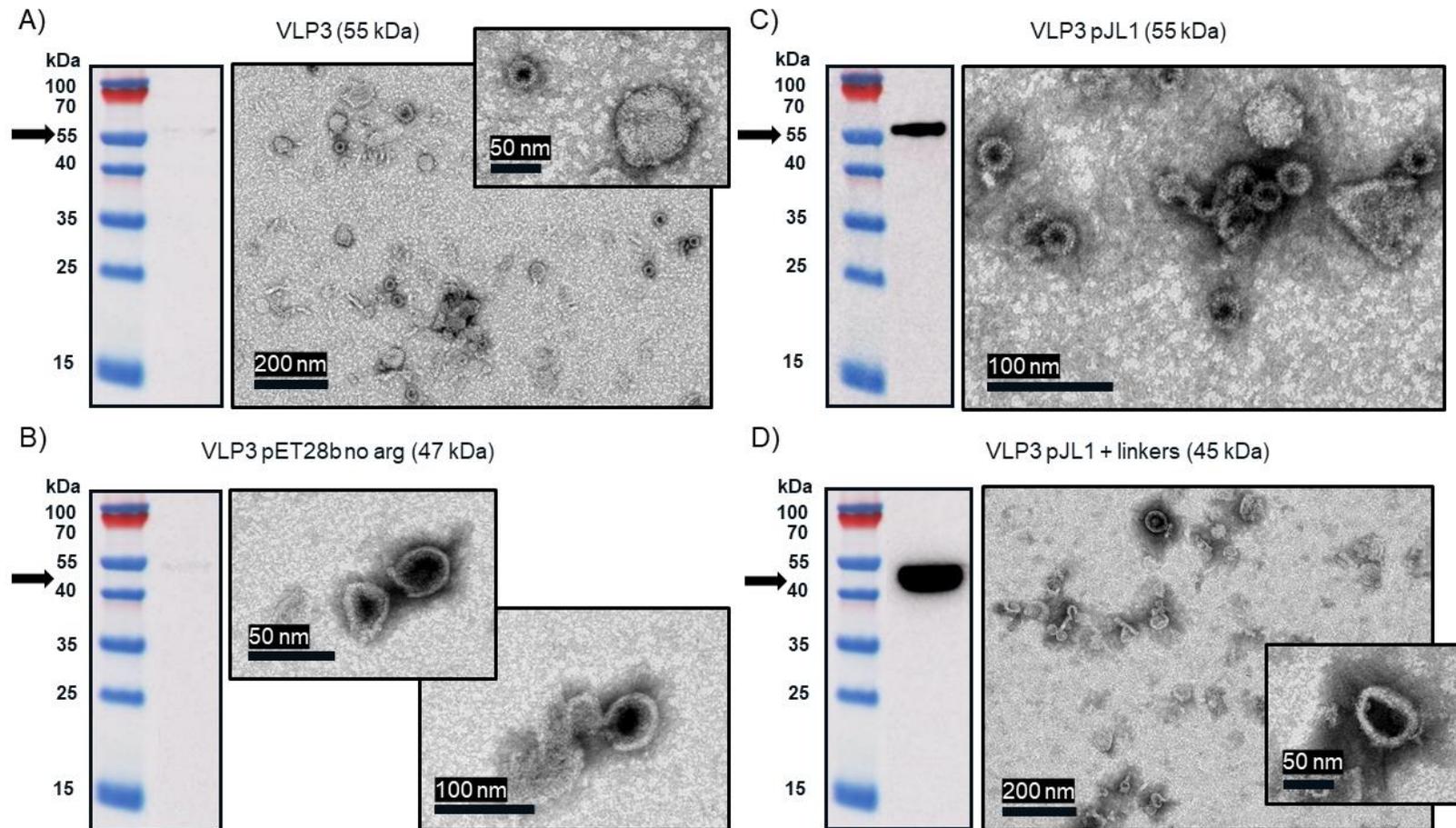


Figure 6.10 Ammonium Sulphate Precipitation and TEM images of VLP 3 Variants

CFPS reactions to produce VLP 3 variants were performed at 18°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in renaturing buffer and saved as the insoluble fraction. The soluble fraction was then precipitated with ammonium sulphate. The resuspended precipitate was analysed via western blot with the [14E11] primary antibody and imaged under a TEM. These TEM images were taken by Ian Brown at the University of Kent. The black arrow indicates the full-length product.

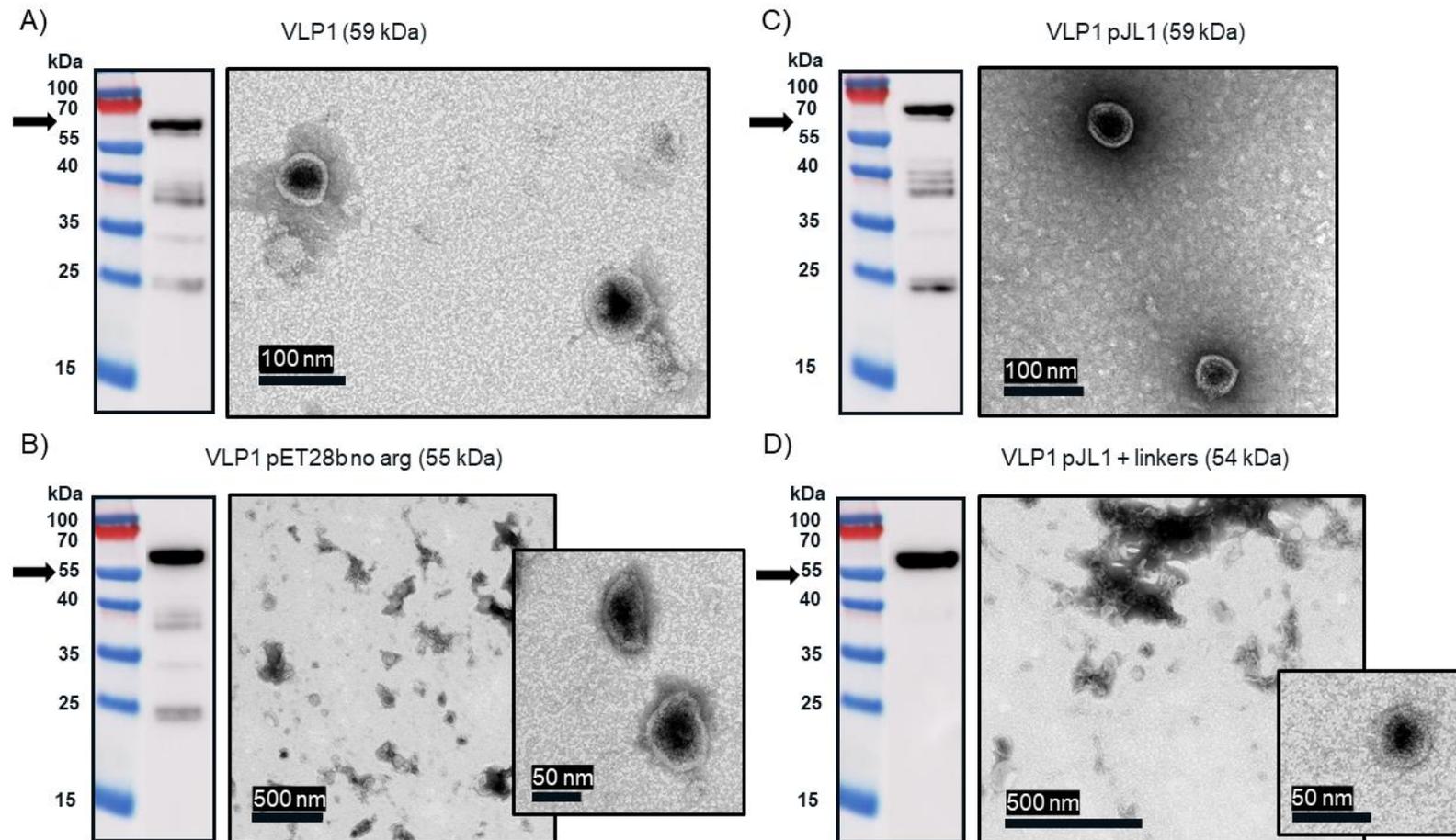


Figure 6.11 Ammonium Sulphate Precipitation and TEM images of VLP 1 Variants

CFPS reactions to produce VLP 1 variants were performed at 30°C. These reactions were centrifuged at 13000 rpm for 10 minutes. The supernatant was decanted and saved as the soluble fraction. The pellet was resuspended in renaturing buffer and saved as the insoluble fraction. The insoluble fraction was then precipitated with ammonium sulphate. The resuspended precipitate was analysed via western blot with the [14E11] primary antibody and imaged under a TEM. These TEM images were taken by Ian Brown at the University of Kent. The black arrow indicates the full-length product.

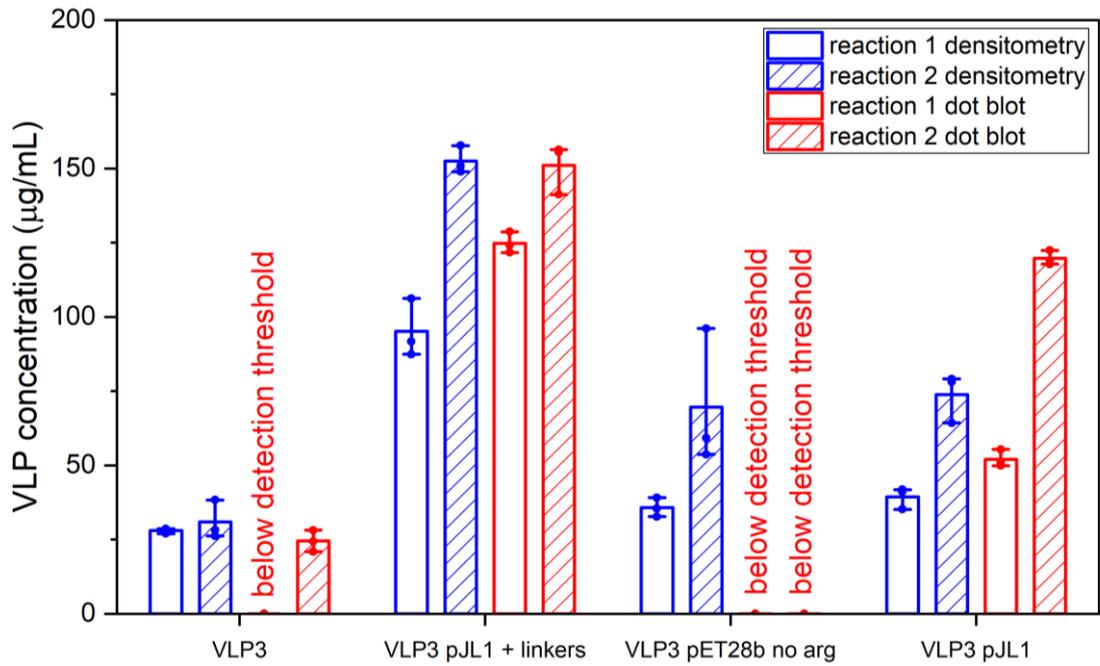


Figure 6.12 Densitometry and Dot Blot Analysis to determine titre for VLP 3 variants

Resuspended precipitate from the CFPS reactions producing VLP 3 variants were analysed using a densitometry reading of an SDS-PAGE analysis and a dot blot analysis for each sample. Two sets of reactions were performed (reaction 1 and reaction 2). Error bars represent plus or minus one standard deviation for n = 3 technical replicates, each represented as a single data point.

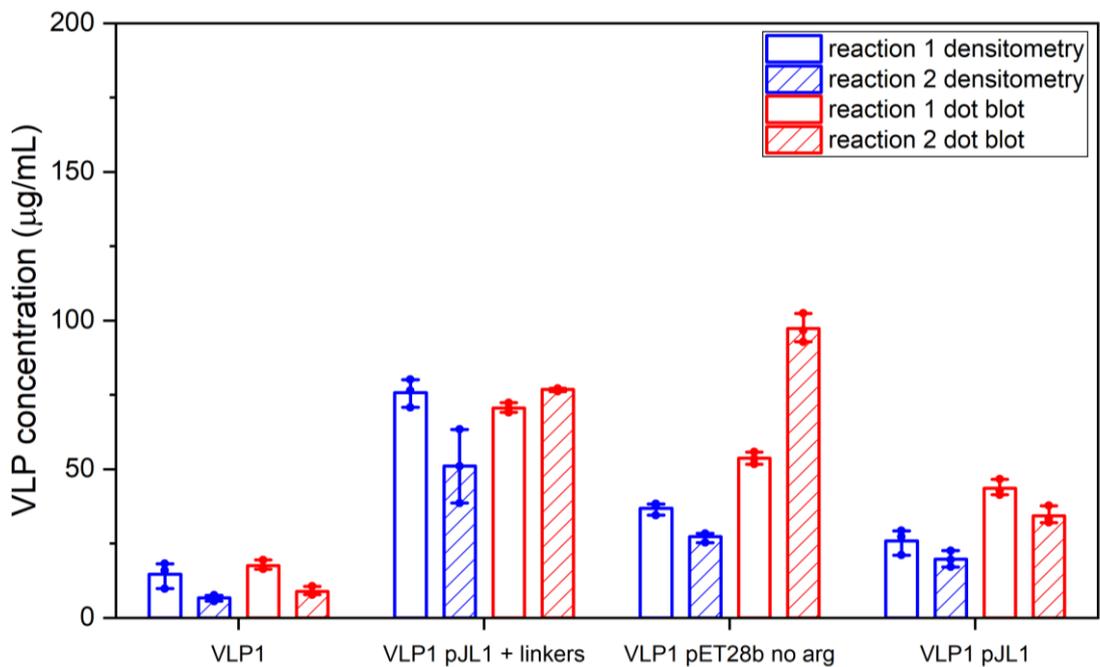


Figure 6.13 Densitometry and Dot Blot Analysis to determine titre for VLP 1 variants

Resuspended precipitate from the CFPS reactions producing VLP 1 variants were analysed using a densitometry reading of an SDS-PAGE analysis and a dot blot analysis for each sample. Two sets of reactions were performed (reaction 1 and reaction 2). Error bars represent plus or minus one standard deviation for n = 3 technical replicates, each represented as a single data point.



Table 6.1 Densitometry and Dot Blot Analysis to determine titre for VLP 3 variants

	<b>Densitometry Reaction 1 (<math>\mu\text{g/mL}</math>)</b>	<b>Densitometry Reaction 2 (<math>\mu\text{g/mL}</math>)</b>	<b>Dot Blot Reaction 1 (<math>\mu\text{g/mL}</math>)</b>	<b>Dot Blot Reaction 2 (<math>\mu\text{g/mL}</math>)</b>
<i>VLP 3</i>	28.0 $\pm$ 0.9	30.9 $\pm$ 6.5	--*	24.5 $\pm$ 3.6
<i>VLP 3 pJL1</i>	39.4 $\pm$ 3.6	73.8 $\pm$ 8.3	52.1 $\pm$ 2.9	119.8 $\pm$ 2.4
<i>VLP 3 pET28b no arg</i>	35.8 $\pm$ 3.2	69.7 $\pm$ 23.1	--*	--*
<i>VLP 3 pJL1 + linkers</i>	95.1 $\pm$ 9.9	152.5 $\pm$ 4.6	124.8 $\pm$ 3.6	151.0 $\pm$ 8.5

\*these samples were below the threshold of detection

Table 6.2 Densitometry and Dot Blot Analysis to determine titre for VLP 1 variants

	<b>Densitometry Reaction 1 (<math>\mu\text{g/mL}</math>)</b>	<b>Densitometry Reaction 2 (<math>\mu\text{g/mL}</math>)</b>	<b>Dot Blot Reaction 1 (<math>\mu\text{g/mL}</math>)</b>	<b>Dot Blot Reaction 2 (<math>\mu\text{g/mL}</math>)</b>
<i>VLP 1</i>	14.6 $\pm$ 4.3	6.7 $\pm$ 1.0	17.6 $\pm$ 1.6	8.9 $\pm$ 1.5
<i>VLP 1 pJL1</i>	25.9 $\pm$ 4.3	19.7 $\pm$ 2.8	43.6 $\pm$ 2.7	34.3 $\pm$ 3.0
<i>VLP 1 pET28b no arg</i>	36.9 $\pm$ 2.1	27.4 $\pm$ 1.8	53.7 $\pm$ 2.0	97.3 $\pm$ 4.8
<i>VLP 1 pJL1 + linkers</i>	75.8 $\pm$ 4.7	51.0 $\pm$ 12.4	70.6 $\pm$ 1.6	76.8 $\pm$ 0.6

## 6.5 Conclusion

Using the CFPS manufacturing system in this project four tandem-core HBcAg VLP constructs were produced, including two universal influenza vaccine candidates, VLP 3 and VLP 1. By decreasing the temperature of the reactions producing VLP 3, the solubility of the product was improved. Decreasing the temperature in the reaction producing VLP 1 did not have the same effect – it remained insoluble. Particles in the expected ~40 nm diameter range were present in reactions producing VLP 3, although some aggregate material was also present. Reactions producing VLP 1 contained particle-like structures in the ~60 nm range that may not have formed properly due to steric effects. While the desired full-length product was successfully produced, lower molecular weight product related impurities were also produced. By comparing reactions with and without protease

inhibitor, examining the variations in western blots performed with different primary antibodies that correspond to different parts of the tandem-core VLP protein, and contrasting the western blot analyses to the accompanying SDS-PAGE analyses, it was determined that these impurities were not the result of protease cleavage. They are either partially transcribed or partially translated species. Further work quantifying mRNA, tRNA, and ribosomes will need to be done to establish whether this is an issue with transcription, translation, or both processes.

Using the CFPS platform in this project, three variants on the two vaccine candidates were quickly expressed and analysed. As an attempt at decreasing the presence of product related impurities and improving titre, assembly, and solubility of the vaccine products, the constructs were modified with a plasmid backbone optimised for CFPS expression, the arginine rich region at the C-terminus of the construct was removed, and the linkers around the influenza A inserts were adjusted. As with the original constructs, the reactions producing VLP 3 variants contained particles in the expected size range and the reactions producing VLP 1 variants contained larger aggregate species. It was determined that using an optimised plasmid backbone improves titres but does not impact lower molecular weight product related impurities. Removing the arginine-rich region may decrease the presence of lower molecular weight product related impurities but results in ill-formed particles. Optimising the backbone, removing the arginine rich region, and adjusting the linkers significantly decreases the presence of lower molecular weight product related impurities and improves titres but does not result in consistently symmetrical particles. By expressing assembled tandem-core HBcAg VLPs in the CFPS system in this project and designing and expressing improved versions of those VLPs, it has been demonstrated that this CFPS system is not only capable of generating complex self-assembling products for vaccine

production, but it is a useful tool for iteration and improvement of vaccine product design.

## 7. Conclusions and Future Work

### 7.1 Review of Project Objectives

The overarching goal of this project was to establish a novel manufacturing system that could be used to produce self-assembling particles for vaccine and gene therapy applications in under a day. In line with achieving this goal, this project had three objectives: 1) develop an on-demand CFPS manufacturing system at UCL, 2) design a process development strategy for this manufacturing system, and 3) use the system to express vaccine and gene therapy products.

A rapid, robust, and flexible *E. coli*-based CFPS platform was developed that consistently produced over 400 µg/mL superfolder green fluorescent protein (sfGFP) in 4 hours. This was achieved by thoroughly examining the three major components of the system: the cell extract, the concentrated reaction mixture, and the plasmid backbone. Five different *E. coli* strains for the cell extract were tested. Three strains, BL21-Star™, ClearColi®, and SHuffle resulted in yields over 450 µg/mL. The impact of IPTG-induction of the cells during growth was also investigated and it was determined that induced extracts did not require additional T7 RNA polymerase to achieve similar titres. Two concentrated reaction mixes were compared, and it was found that a more complex mixture that used glycolysis to regenerate ATP (the PANoxSP system) outperformed a more minimal mixture that used oxidative phosphorylation (the Cytomim system). While any plasmid with the gene of interest under the T7 promoter can be used in this CFPS system, shorter plasmid backbones optimised for CFPS resulted in higher titres than plasmid backbones designed for *in vivo* production. It was also demonstrated that this system could be scaled up by volume over three orders of magnitude in deep well flower plates, shake flasks, T-75 flasks, and small-scale bioreactors.

Next, a process development strategy for this system was designed and it was validated with two model proteins, sfGFP and hepatitis B core antigen (HBcAg). sfGFP production was improved by 38% and HBcAg production by 190%. This was achieved by investigating the impact on the reaction titre of the concentration of the plasmid, the amount of the cell extract, the pH of the concentrated reaction mix, the temperature of the reaction, and the length of the reaction. Using multivariate data analysis (MVDA), the parameters that were most critical for each product were determined. Then this model was validated using design of experiments (DoE) to increase titre. These findings resulted in the formation of a three-step process development strategy that can be completed in as little as 48 hours.

Finally, this CFPS system was used to produce vaccine and gene therapy products to mimic personalised medicines. The three capsid proteins from adeno-associated virus serotype 2, which can assemble to form AAV2 VLPs, were expressed. These VLPs can deliver genetic material and may be useful in the development of gene therapy products. All three capsid proteins, VP1, VP2, and VP3 with and without a 6x histidine tag were produced. The VLPs were not assembled in the CFPS reactions, but the proteins were purified with the 6x histidine tag using a Ni-NTA resin, although very little product remained in the elution fraction. Once the recovery from the resin is improved, the proteins can be combined, and the capsids can be assembled *in vitro*. This work may pave the way for CFPS as a platform for gene therapy and personalised medicines manufacturing.

The tandem-core HBcAg VLPs developed as universal influenza vaccine candidates were also expressed because they self-assemble, they contain proteins from two different parent viruses, and they tend to aggregate when

expressed *in vivo*. The production of VLP 3, a tandem-core HBcAg VLP with three versions of the matrix 2 ectodomain (M2e) protein in its major insertion regions (MIRs), and VLP 1, which displays the haemagglutinin stalk protein (HA2) in one its MIRs and three versions of the M2e protein in the other was central to this project. While full-length product was produced for both tandem-core vaccine candidates, several lower molecular weight product related impurities were also produced. After determining that these impurities were not the result of proteases, the CFPS system was used to rapidly test new plasmid designs for both products. Three modified constructs were examined, one in which the original gene sequence was in a plasmid backbone optimised for CFPS, one in the original plasmid backbone with a gene sequence in which the arginine-rich region was removed to improve stability, and one in the plasmid backbone optimised for CFPS with the arginine-rich region removed and with more uniform peptide linkers around the influenza A inserts. It was determined that using the optimised plasmid backbone improved titres by 1.8 times over the original VLP 1 construct and 1.4 times over the original VLP 3 construct and that titres were further increased with constructs in which the linker lengths were more consistent. However, any construct with the arginine-rich region removed did not result in consistently symmetrical particles. This work has shown that the CFPS system in this project can be used to rapidly express different modified constructs to improve the design of a biotherapeutic or vaccine.

In summary, all three objectives have been achieved. While there are no doubt improvements that could be made to this system, this project has established a foundation for an *E. coli*-based CFPS system for on-demand production of personalised medicines. This will allow for a transition away from the traditional

“one-size-fits-all” mass production model to better accommodate novel stratified and personalised medicines.

## **7.2 Recommendations for Future Work**

This doctoral project is part of the Future Targeted Healthcare Manufacturing (FTHM) Hub which was founded to investigate the manufacturing, business, and regulatory challenges that would need to be overcome in order to develop and produce new targeted biologic medicines in a quick and cost-effective manner. If this project were to be taken forward, then the challenges of using an *E. coli*-based CFPS system for on-demand production of personalised medicines should also be examined from manufacturing, business, and regulatory perspectives and improved accordingly (Figure 7.1).

Manufacturing	Business	Regulatory
Use the CFPS system to generate a variety of products	Generate T7 RNA polymerase in-house	Maintain SOPs
Develop a traditional chromatography set-up for CFPS reactions	Develop a cost-effective concentrated reaction mixture using maltose	Optimise cell growth and lysis conditions for extract preparation
Integrate chromatography resin directly in CFPS reactions	Label plasmids with magnetic beads for recycling	Implement PCR clean-up for plasmid preparation
Design a personalised medicines product (delivery or display mechanism) and express it via CFPS	Using continuous culture to generate cells for extract preparation	Implement consistency tests for concentrated reaction mix, amino acid solutions, and T7 RNA polymerase
Address challenges with self-assembling particles (see Figure 7.2)	Optimise concentrated reaction mixture to significantly reduce costs	Use RT-qPCR, RNA-seq, and ribosome labelling to understand transcriptional and translational limitations
Develop an understanding of the transcriptional and translational limitations of the CFPS proteasome	Establish separate extract production and CFPS expression companies	Fully automate CFPS reactions
		Develop inline monitoring of process parameters

Relatively easy (<1 year)

Challenging (1-5 years)

Difficult (>5 years)

Figure 7.1 Manufacturing, Business, and Regulatory Challenges for CFPS Production of Personalised Medicines

### 7.2.1 Opportunities for CFPS in Biomanufacturing

Before thoroughly examining the manufacturing, business, and regulatory challenges of an on-demand *E. coli*-based CFPS production system, it is important to acknowledge CFPS in the biopharmaceutical manufacturing space as a whole. Although CFPS first emerged in the 1960s, it has only been recognised as a system for biopharmaceutical manufacturing in the last decade. While it has its advantages, namely, quick and flexible reactions generating high titres, it also expensive, under-developed and a regulatory enigma. It is extremely unlikely that CFPS reactions will become common production processes any time soon. Nor should they. The traditional manufacturing approach has been studied and optimised for decades to mass produce biologics. There is no way that CFPS process could conceivably replace those manufacturing schemes. CFPS is better suited to small scale, rapid production processes. Therefore, CFPS would be beneficial for the production of personalised medicines where very few doses need to be created. Quick and mobile production processes might also be helpful in disaster scenarios where the cold chain has been interrupted or in pandemic response; this was the intention behind the Bio-MOD device that uses a mammalian CFPS system (Adiga et al., 2018). Another potential use of this system is for clinical grade production material for characterisation studies or rapid screening of new pharmaceutical candidates, an application Swiftscale Biologics is just now bringing to the market (Biologics, 2020). As demonstrated in Chapters 4 and 6, it is an excellent screening tool. A large variety of process conditions or molecule designs can be tested in a matter of hours, helping to shorten product development and process development timelines. CFPS is also a strong candidate for processes that are difficult *in vivo* like the production of membrane proteins, toxic proteins, and proteins with non-standard amino acids. The system developed

in this project was used to test a few such difficult-to-express proteins (see Appendix B). These applications may become even more widespread as eukaryotic CFPS systems are further developed, potentially allowing for a variety of post-translational modifications.

### 7.2.2 Addressing Manufacturing Challenges

In this project, it was determined that an *E. coli*-based CFPS system could be used to rapidly express self-assembling products. However, this CFPS system is still a long way from being a complete on-demand personalised medicines manufacturing scheme. Two additional goals would need to be accomplished for that to happen: 1) use the CFPS system to express a personalised medicine product, and 2) incorporate the CFPS reaction (the upstream production process) with downstream purification processes for a complete on-demand process.

The first goal could be achieved by expressing a personalised medicine product developed by another group. This sounds easy enough, but the product would need to be susceptible to the *E. coli*-based CFPS environment. Enveloped products or products that require advanced post-translational modifications would not be possible, although the system could be adjusted with IAM and DsbC for disulphide bond formation (Knapp et al., 2006). The first goal could also be achieved by designing a novel personalised medicine based on knowledge of self-assembling systems obtained in this project. This novel product may be a delivery mechanism, similar to current viral vector therapies, that carries cargo to be delivered or it might be a display mechanism, where patient-specific proteins are presented on the outside of the particle. For a personalised medicine operating as a delivery mechanism, one potential candidate is the AAV2 capsid proteins expressed in Chapter 5. With an improved purification scheme, they could be isolated and then assembled around genetic material, perhaps a copy of the

*RPE65* gene for the treatment of retinal dystrophy like Luxturna® or the *SMN1* gene for the treatment of spinal muscular atrophy like Zolgensma®. A potential personalised medicines candidate operating as a display mechanism could be a VLP product or an encapsulin derived-product that displays patient-specific antigens much like the tandem-core HBcAg VLPs display influenza A antigens in their MIRs (Rohovie et al., 2017, Diaz et al., 2020). These therapies may even be able to be modular, where the same self-assembling particle is used as a base to display a variety of other antigens that can be attached using click chemistry (Patel and Swartz, 2011).

The second goal related to manufacturing, incorporating upstream production processes with downstream purification processes for a complete on-demand process, could be achieved by designing a system with a fully integrated process like the one in the InSCyT system (Crowell et al., 2018). If ion-exchange chromatography is part of the system, this would require a robust understanding of the system's proteasome and metabolism as well as the development of models to determine the proper column and buffer conditions for a given product. Luckily, some groups are already examining these areas and, because the same extract from the same bacterial strain could be used for multiple products, this work may result in a series of highly similar purification processes for a variety of products (Foshag et al., 2018, Miguez et al., 2019, Vecchiarello et al., 2019). The better the understanding of the system is, the more flexibility can be engineered into it. Chromatography with a resin column is currently the most common purification scheme used for large-scale virus and virus-like particle preparation, but alternative processes may come forward in the future, like monolith columns (Burden et al., 2012). Alternatively, because the CFPS reactions are open to the environment, products could be bound directly to the chromatography resin during

the reaction and then eluted to eliminate additional downstream processing steps. Once an appropriate downstream purification scheme is determined, more rigorous testing can be performed on the product of interest. Endotoxin levels could be analysed using a Limulus Amebocyte Lysate (LAL) assay to ensure minimal lipopolysaccharide (LPS) is present or nucleic acid content could be determined measuring the  $\lambda=260/280$  nm absorbance ratio.

#### *7.2.2.1 Addressing Manufacturing Challenges Specific to Self-Assembling*

##### *Particles*

While CFPS reactions have many advantages when it comes to self-assembling particle production, like reactions that yield high titres in short time frames for rapid screening and open reactions that allowed for non-physiological reaction conditions for improved assembly, all expression systems for self-assembling particles must address the challenges that arise when determining the appropriate methods for solubilising, purifying, quantifying, and characterising these particles. For tandem-core HBcAg VLPs, the lack of established processes makes these challenges even more difficult. Significant work would need to be done on all four processes if they are going to continue to be the model for this system. It is for that reason that further work with these particular vaccine candidates is not recommended. It would be better to use a product that has already undergone extensive solubilisation, purification, quantification, and characterisation work, like the human papilloma virus VLP, the hepatitis B surface antigen VLP, the norovirus VLP or other modified hepatitis B core antigen VLPs that have fewer inserts and lack a 6x histidine tag.

The 6x histidine tag contributes significantly to the solubilisation and purification issues with this product. When the particles assemble, the 6x histidine tag is not displayed on the outside so that it might be used for Ni-NTA column

chromatography. Instead, it sits inside the particle and interacts with other 6x histidine tag to stabilise the particle, making it nearly impossible to disassemble the particles and insoluble aggregates for purification as monomers (Schumacher et al., 2018). To better solubilise these proteins, the 6x histidine tag could be removed and larger peptide tags like maltose-binding protein (MBP), glutathione S-transferase (GST), thioredoxin (TrxA), DsbC, small ubiquitin-like modifier protein (SUMO), or N-utilisation substance A (NusA) could be incorporated (García-Fruitós, 2015). This would prevent the assembly of the particles which may allow for purification using affinity chromatography. However, the tags would then need to be removed and the correct buffer conditions that result in maximum proper VLP assembly and minimum VLP-like aggregates would need to be determined. Additional steps like these are costly and may be difficult to incorporate into an on-demand system.

The VLPs could be purified using multiple column chromatography steps including affinity chromatography (if the tags are used as suggested above), SEC, IEX, and HIC (Ramirez et al., 2018). This would require the generation of a significant amount of VLP material to run optimisation studies. Because yields are relatively low, that material may need to be concentrated either before or after being run on a chromatography column. Several membranes were used throughout this project including the Vivaspin® Centrifugal Concentrators and 0.22 µm Costar® Spin-x® centrifuge tube filters. Each time a substantial amount of material was lost (data not shown). It appears that the VLPs or VLP-like aggregates get stuck to the membranes. Several membranes would need to be tested before the VLPs could be concentrated effectively.

Dot blot analysis and SDS-PAGE densitometry analysis were used to quantify tandem-core HBcAg VLP titres. An ELISA would be a more accurate

method of quantification. Unfortunately, many of the ELISA kits for HBcAg detection that are commercially available use a primary antibody that binds to amino acids 70-80 in the HBcAg monomer. The tandem-core HBcAg constructs have been modified beginning at amino acid 77; that is where the influenza A inserts are located. Fortunately, there are other primary antibodies available that bind to different regions of HBcAg. Because the VLPs are assembled (or in some cases, partially assembled or aggregated), that may impact whether the antibodies are able to bind in certain areas. Several different antibodies would need to be tested for the creation of an in-house ELISA to most accurately determine the titre of the tandem-core HBcAg VLPs.

Characterisation of the VLPs was done by TEM in this project. That would not be feasible in an on-demand manufacturing system. Miniaturised versions of processes like DLS or asymmetric flow field-flow fractionation (AF4) might be used in the future if the material is homogenous enough after purification (Chaun et al., 2008). If the tandem-core HBcAg VLP vaccine candidates were a transformative one-of-a-kind live-saving product that could not be made any other way, then perhaps it would be worth it to invest the money, time, and talent that would be needed to optimise these processes and produce these vaccine candidates. As it is, there are a variety of other universal influenza vaccine products that are further along in the development process. Twenty-two technologies have reached clinical trials, three of which are in Phase 3 clinical trials (Ostrowsky et al., 2020).

#### *7.2.2.2 Suggestions for Continued Work with Tandem-Core HBcAg VLPs and AAV2 VLPs*

Although several reasons why it will be difficult to work with the tandem-core HBcAg constructs going forward are detailed in the previous section, if another researcher were to undertake a continuation of the work in this thesis, I have a few

suggestions on where they might begin. I would recommend generating one more iteration of VLP 1 and VLP 3. These modified constructs would be the original sequence with the modified linkers (as in VLP 1 pJL1 + linkers and VLP 3 pJL1 + linkers) and the arginine-rich region in the pJL1 backbone. Based on the results in Chapter 6, I would anticipate that particle formation would be more symmetrical because the arginine-rich region is still present, but that there may be fewer lower molecular weight product related impurities because of the adjustments made to the linkers. If that does not prove to be true with the expression of these proposed constructs, then additional iterations on the vaccine candidates could be investigated. At that point, it is probably in the researcher's best interest to remove the 6x histidine tag from the construct or replace the 6x histidine tag with another tag or protein. That may allow for the vaccine candidates to be purified as monomers and then assembled separately. It would also be wise to work primarily with VLP 3. In order to overcome the steric effects preventing proper assembly of VLP 1, the influenza inserts will need to be dramatically changed or potentially removed and until those issues are mitigated, VLP 1 will not assemble properly.

For the AAV2 capsid proteins expressed in Chapter 5, I recommend developing a purification strategy for the capsid proteins. The capsid proteins with the 6x histidine tag bind to Ni-NTA resin in microcentrifuge tubes; this process could be transferred to a gravity flow column or an FPLC. Because the 6x histidine tags may impede capsid assembly, a purification strategy for the capsid proteins without the tag will also be necessary. As they are insoluble, the capsid proteins may be able to be precipitated, although as observed with the tandem-core HBcAg vaccine candidates, that will also precipitate several proteins present in the CFPS reaction. A variety of chromatography techniques including affinity chromatography, IEX, and HIC have been used to purify AAV2 viral vectors

(Adams et al., 2020). Perhaps one of these systems could be adapted for the purification of the AAV2 capsid proteins. It might also be beneficial to investigate other tags, particularly tags that improve solubility of the protein as the capsid proteins are largely insoluble. Once the capsid proteins can be purified, the protocol from Le et al. 2018 could be applied to better solubilise them and assemble them into particles (Le et al., 2019). If AAV proteins produced via CFPS can be assembled into VLPs, that would be an important contribution to this field of research as similar multi-protein assemblies have not yet been generated using CFPS systems.

For both the tandem-core HBcAg vaccine candidates and the AAV2 capsid proteins, once the product is in a state where it can be produced consistently and forms symmetrical particles, I would recommend applying the process development strategy from Chapter 4 to improve titres.

Solubilisation	Purification	Quantification	Characterisation
<p data-bbox="159 363 609 464">Replace 6x histidine tag with TrxA, DsbC, SUMO, or NusA</p>	<p data-bbox="647 363 1102 464">Prevent assembly of individual dimers into VLPs</p> <p data-bbox="647 480 1102 651">Design and express a construct with the arginine-region and the adjusted linkers in the pJL1 backbone</p> <p data-bbox="647 671 1102 778">Develop an assembly protocol</p> <p data-bbox="647 799 1102 906">Design a method of concentrating particles</p> <p data-bbox="647 927 1102 1066">Develop a column chromatography protocol (IEX, SEC, or HIC)</p> <p data-bbox="647 1086 1102 1193">Transfer Ni-NTA process to gravity flow or FPLC</p>	<p data-bbox="1137 363 1592 507">Design an ELISA using a compatible anti-HBcAg antibody</p> <p data-bbox="1137 523 1592 667">Improve titres using the process development strategy</p>	<p data-bbox="1628 363 2083 539">Use Dynamic Light Scattering or Asymmetric Flow Field-Flow fractionation to characterize particles</p>

Suggestions for tandem-core HBcAg VLPs
  Suggestions for AAV2 VLPs
  Suggestions for both products

Figure 7.2 Challenges Specific to Self-Assembling Particle Production

### 7.2.3 Addressing Business Challenges

From a business perspective, for CFPS to be a lucrative manufacturing scheme in comparison to a traditional CHO manufacturing process, the cost of goods per reaction would need to be cut in half (Stamatis, 2020). There are already several changes verified by other research groups that could be used to drive down the cost of the CFPS. One of the most common is generating T7 RNA polymerase in-house rather than purchasing it. This could be done by transforming a plasmid expressing T7 RNA polymerase into an *E. coli* expression strain, purifying it, and adding it into the reaction (Failmezger et al., 2017). The plasmid for the T7 RNA polymerase could also be added directly to the CFPS reaction, although co-expression may be dependent on the product being expressed and may not be as consistent as adding in a set amount of T7 RNA polymerase (Caschera and Noireaux, 2015b). Another strategy would be to reduce the cost of the concentrated reaction mixture by adding in maltose to sustain the reaction and prevent the accumulation of inhibiting inorganic phosphate (Caschera and Noireaux, 2014). Alternative energy sources that are not as expensive as PEP could also be examined. DNA plasmid preparation is another a time consuming and expensive process necessary for CFPS reactions. It might be possible to recycle plasmids by labelling them with magnetic beads (Lee et al., 2012). Costs could also be decreased by finding ways to reduce the number of unit operations needed. For example, endotoxin removal is a necessary step in manufacturing schemes that employ microbial fermentation. When endotoxins, most notably lipopolysaccharide (LPS), are present even in small amount in the human blood stream, they illicit a strong innate immune response; at higher concentrations they can result in septic shock (Sampath, 2018). By using strains like ClearColi® (DE3) that have a modified LPS to prevent the formation of endotoxins to begin with, that step can

be circumvented. Using continuous culture to produce cell extracts could bring down costs and reduce batch-to-batch variation as well. Costs might also be reduced by outsourcing the extract preparation process so that one company was producing extract for several companies that were generating products via CFPS, ultimately allowing all companies involved to capitalise on economies of scale (Melinek et al., 2020).

#### 7.2.4 Addressing Regulatory Challenges

There is still a great deal of ambiguity around what will be required of a CFPS manufacturing scheme for it to meet regulatory guidelines. As Sutro Biopharma Inc. accelerates their products into clinical trials, the desired specifications will hopefully become clear to others operating in the CFPS space. Regardless of the exact specifications, improving consistency and quality of material will be advantageous. First and foremost, robust and regularly updated standard operating procedures need to be developed for all steps of the process, including cell extract preparation, concentrated reaction mixture preparation, amino acid solutions preparation, T7 RNA polymerase preparation, and DNA plasmid preparation. Other groups have optimised the length of cell growth and length of induction to improve extract quality (Dopp and Reuel, 2018). Examining the lysis and centrifugation conditions to remove the membranes and the large proteins that tend to aggregate with self-assembling particles might also result in improved cell extracts. Additionally, there have been recommendations on how best to prepare the amino acids solutions and concentrated reaction mixture for consistent performance (Caschera and Noireaux, 2015b, Dopp et al., 2019a). Based on the work with the HBcAg protein in a pET-Duet-1 backbone performed in this project, optimising the concentration of IPTG in the reaction may also help to boost titres when certain plasmid backbones are being used. Introducing PCR

clean up kits alongside plasmid preparation kits would result in DNA plasmids of higher purity (Strychalski and Romantseva, 2020). In addition to improving reagent consistency, improving process consistency by automating the reaction using a robotic system like the Tecan liquid handling system could reduce variability and operating time (Quast et al., 2015, Caschera et al., 2011).

Consistency will also result from a better understanding of the biological processes happening in CFPS systems. At present, the consistency of the cell extract is tested by using that extract to generate sfGFP and comparing the titres to previous reactions with other batches of extract. Similar tests could be done on new batches of concentrated reaction mixture, amino acids solutions, and T7 RNA polymerase. Ideally, the limiting factors in the system would be known and whether they are related to transcription, translation, or are product dependent could be determined. There are techniques available now that could help expand this knowledge: RT-qPCR to quantify mRNA, RNA-Seq to determine the length of mRNA transcripts, or ribosome labelling to better understand where the ribosomes are in the system. However, these processes require analysis after the reaction has already been completed. It would be much more informative if these components as well as tRNAs, elongation factors, initiation factors, and polymerases could be monitored inline. If that were possible, one could imagine a system where one of the essential components in transcription or translation is depleted over the course of the reaction and then added in later on – a sort of hybrid system between the crude lysate and PURE platforms. This could also allow for the codon-optimisation of gene sequences for this particular CFPS system and its limitations.

## 8. References

2011. Expressway™ Cell-Free E. coli Expression System. Available: [https://assets.thermofisher.com/TFS-Assets/LSG/manuals/expressway\\_system\\_man.pdf](https://assets.thermofisher.com/TFS-Assets/LSG/manuals/expressway_system_man.pdf) [Accessed September 27].
2014. Chapter Seven - Salting out of Proteins Using Ammonium Sulfate Precipitation. *In: LORSCH, J. (ed.) Laboratory Methods in Enzymology: Protein Part C*. Baltimore, MD: Elsevier Inc.
- ACHARYA, S. K. & BATRA, Y. 2005. Hepatitis B Virus Genotypes: Clinical Implications. *Association of Physician of India Medicine Update*, 87, 425-429.
- ADAMS, B., BAK, H. & TUSTIAN, A. D. 2020. Moving from the bench towards a large scale, industrial platform process for adeno-associated viral vector purification. *Biotechnology and Bioengineering*, 117, 3199-3211.
- ADIGA, R., AL-ADHAMI, M., ANDAR, A., BORHANI, S., BROWN, S., BURGENSEN, D., COOPER, M. A., DELDARI, S., FREY, D. D., GE, X., GUO, H., GURRAMKONDA, C., JENSEN, P., KOSTOV, Y., LACOURSE, W., LIU, Y., MOREIRA, A., MUPPARAPU, K., PEÑALBER-JOHNSTONE, C., PILLI, M., PUNSHON-SMITH, B., RAO, A., RAO, G., RAUNIYAR, P., SNOVIDA, S., TAURANI, K., TILAHUN, D., TOLOSA, L., TOLOSA, M., TRAN, K., VATTEM, K., VEERARAGHAVAN, S., WAGNER, B., WILHIDE, J., WOOD, D. W. & ZUBER, A. 2018. Point-of-care production of therapeutic proteins of good-manufacturing-practice quality. *Nature Biomedical Engineering*, 2, 675-686.
- AHN, J.-H., HWANG, M.-Y., OH, I.-S., PARK, K.-M., HAHN, G.-H., CHOI, C.-Y. & KIM, D.-M. 2006. Preparation method for *Escherichia coli* S30 extracts completely dependent upon tRNA addition to catalyze cell-free protein synthesis. *Biotechnology and Bioprocess Engineering*, 11, 420.
- ANDAR, A. U., DELDARI, S., GUTIERREZ, E., BURGENSEN, D., AL-ADHAMI, M., GURRAMKONDA, C., TOLOSA, L., KOSTOV, Y., FREY, D. D. & RAO, G. 2019. Low-cost customizable microscale toolkit for rapid screening and purification of therapeutic proteins. *Biotechnology and Bioengineering*, 116, 870-881.
- AVEXIS. 2020. *Zolgensma targets the genetic root cause of SMA* [Online]. Available: [https://www.zolgensma.com/how-zolgensma-works?gclid=CjwKCAjw1v\\_0BRakEiwALFki5pt9PoCYD0X0P4TD9pK17Hf41230Yj-5a7ExPx4J\\_LDrF\\_F94y0dLxoCz4sQAvD\\_BwE](https://www.zolgensma.com/how-zolgensma-works?gclid=CjwKCAjw1v_0BRakEiwALFki5pt9PoCYD0X0P4TD9pK17Hf41230Yj-5a7ExPx4J_LDrF_F94y0dLxoCz4sQAvD_BwE) [Accessed].
- BACHMANN, M. F. & JENNINGS, G. T. 2004. Virus-like particles: Combining innate and adaptive immunity for effective vaccination. *In: KAUFMANN, S. H. E. (ed.) Novel vaccinatino strategies*. Weinheim: Wiley-VCH Verlag GmbH & Co. KGaA.

- BACHMANN, M. F., ROHRER, U. H., KUNDIG, T. M., BURKI, K., HENGARTNER, H. & ZINKERNAGEL, R. M. 1993. The influence of antigen organization on B cell responsiveness. *Science*, 262, 1448-1451.
- BACKOVIC, A., CERVELLI, T., SALVETTI, A., ZENTILIN, L., GIACCA, M. & GALLI, A. 2012. Capsid protein expression and adeno-associated virus like particles assembly in *Saccharomyces cerevisiae*. *Microbial Cell Factories*, 11, 124.
- BECERRA, S. P., KOCZOT, F., FABISCH, P. & ROSE, J. A. 1988. Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. *Journal of virology*, 62, 2745-2754.
- BIOLOGICS, S. 2020. *Speed COVID-19 Treatments to Patients with Rapid Antibody Manufacturing* [Online]. Available: <https://www.swiftcovid.com/> [Accessed 2020].
- BIOPHARMA, S. 2019. Available: <https://www.sutro.bio.com/pipeline/> [Accessed].
- BIOSCIENCES, G. 2020. Available: <https://www.greenlightbiosciences.com/markets/plant-sciences/> [Accessed].
- BIOSCIENCES, V. 2019. Three is the magic number in gene therapy production. Nature Research.
- BIOSTRUCTURE, C. 2020. *Virus-Like Particles (VLPs)* [Online]. Available: <https://www.creative-biostructure.com/pclass-virus-like-particles-vlps-1.htm> [Accessed].
- BIRD, L. E., RADA, H., VERMA, A., GASPER, R., BIRCH, J., JENNIONS, M., LOWE, J., MORAES, I. & OWENS, R. J. 2015. Green Fluorescent Protein-based Expression screening of Membrane Proteins in *Escherichia coli*. *Journal of Visualized Experiments* 95.
- BJERGA, G. E. K., LALE, R. & WILLIAMSON, A. K. 2016. Engineering low-temperature expression systems for heterologous production of cold-adapted enzymes. *Bioengineered*, 7, 33-38.
- BLAHA, B. A. F. 2019. *Platform processes for vaccine production: Development of a universal influenza vaccine using Pichia pastoris*. Doctorate of Philosophy, University College London.
- BLESSING, D., VACHEY, G., PYTHOUD, C., REY, M., PADRUN, V., WURM, F. M., SCHNEIDER, B. L. & DÉGLON, N. 2018. Scalable Production of AAV Vectors in Orbitally Shaken HEK293 Cells. *Molecular therapy. Methods & clinical development*, 13, 14-26.
- BOIGARD, H., ALIMOVA, A., MARTIN, G. R., KATZ, A., GOTTLIEB, P. & GALARZA, J. M. 2017. Zika virus-like particle ( VLP ) based vaccine. *PLOS Neglected Tropical Disease*, 11, 1-20.

- BREMER, H. & DENNIS, P. P. 2008. Modulation of Chemical Composition and Other Parameters of the Cell at Different Exponential Growth Rates. *Ecosal Plus*.
- BUNDY, B. C., FRANCISZKOWICZ, M. J. & SWARTZ, J. R. 2007. Escherichia coli-based cell-free synthesis of virus-like particles. *Biotechnology and Bioengineering*, 100, 28–37.
- BUNDY, B. C. & SWARTZ, J. R. 2011. Efficient disulfide bond formation in virus-like particles. *Journal of biotechnology*, 154, 230-239.
- BURDEN, C. S., JIN, J., PODGORNIK, A. & BRACEWELL, D. G. 2012. A monolith purification process for virus-like particles from yeast homogenate. *Journal of Chromatography B*, 880, 82-89.
- CAI, Q., HANSON, J. A., STEINER, A. R., TRAN, C., MASIKAT, M. R., CHEN, R., ZAWADA, J. F., SATO, A. K., HALLAM, T. J. & YIN, G. 2015. A simplified and robust protocol for immunoglobulin expression in Escherichia coli cell-free protein synthesis systems. *Biotechnology Progress*, 31, 823-831.
- CARLSON, E. D., GAN, R., HODGMAN, E. C. & JEWETT, M. C. 2014. Cell-Free Protein Synthesis : Applications Come of Age. *Biotechnology Advances*, 30, 1185-1194.
- CARTER, C., HOUSER, K. V., YAMSHCHIKOV, G. V., BELLAMY, A. R., MAY, J., ENAMA, M. E., SARWAR, U., LARKIN, B., BAILER, R. T. & KOUP, R. 2019. Safety and immunogenicity of investigational seasonal influenza hemagglutinin DNA vaccine followed by trivalent inactivated vaccine administered intradermally or intramuscularly in healthy adults: An open-label randomized phase 1 clinical trial. *PloS one*, 14.
- CARTER, N. J. & CURRAN, M. P. 2011. Live Attenuated Influenza Vaccine (FluMist®; Fluenz™). *Drugs*, 71, 1591-1622.
- CASCHERA, F., BEDAU, M. A., BUCHANAN, A., CAWSE, J., DE LUCREZIA, D., GAZZOLA, G., HANCZYC, M. M. & PACKARD, N. H. 2011. Coping with complexity: Machine learning optimization of cell-free protein synthesis. *Biotechnology and Bioengineering*, 108, 2218-2228.
- CASCHERA, F. & NOIREAUX, V. 2014. Synthesis of 2.3 mg/mL of protein with an all Escherichia coli cell-free transcription-translation system. *Biochimie*, 99, 162-168.
- CASCHERA, F. & NOIREAUX, V. 2015a. A cost-effective polyphosphate-based metabolism fuels an all E. coli cell-free expression system. *Metabolic Engineering*, 27, 29-37.
- CASCHERA, F. & NOIREAUX, V. 2015b. Preparation of amino acid mixtures for cell-free expression systems. *BioTechniques*, 58, 40-43.
- CASTAN, A., SCHULZ, P., WENGER, T. & FISCHER, S. 2018. Cell Line Development. *Biopharmaceutical Processing: Development, Design, and*

*Implementation of Manufacturing Processes*. Amsterdam, Netherlands: Elsevier.

- CATAPULT, C. A. G. T. 2019. Cell and Gene Therapy Catapult Clinical Trials Database 2019. London.
- CERES, P. & ZLOTNICK, A. 2002. Weak Protein-Protein Interactions Are Sufficient for Efficient Assembly of Hepatitis B Virus Capsids. *Biochemistry*, 41, 11525-11531.
- CHACKERIAN, B. 2010. Virus-like particle based vaccines for Alzheimer disease. *Human Vaccines*, 6, 962-930.
- CHACKERIAN, B., LOWRY, D. R. & SCHILLER, J. T. 2001. Conjugation of a self-antigen to papillomavirus-like particles allows for efficient induction of protective autoantibodies. *The Journal of Clinical Investigation*, 108, 415-423.
- CHAO, C. N., LIN, M. C., FANG, C. Y., CHEN, P. L., CHANG, D., SHEN, C. H. & WU, M. 2016. Gene Therapy for Human Lung Adenocarcinoma Using a Suicide Gene Driven by Lung-Specific Promoter Delivered by JC Virus-Like Particles. *PLoS ONE*, 11.
- CHAUN, Y. P., FAN, Y. Y., LUA, L. & MIDDELBERG, A. P. J. 2008. Quantitative analysis of virus-like particle size and distribution by field-flow fractionation. *Biotechnology and Bioengineering*, 99, 1425-1433.
- CHIZZOLINI, F., FORLIN, M., MARTIN, N. Y., BERLOFFA, G., CECCHI, D. & MANSY, S. S. 2017. Cell-Free Translation Is More Variable than Transcription. *ACS Synthetic Biology*, 6, 638-647.
- CHUMPOLKULWONG, N., HORI-TAKEMOTO, C., HOSAKA, T., INAOKA, T., KIGAWA, T., SHIROUZU, M., OCHI, K. & YOKOYAMA, S. 2004. Effects of Escherichia coli ribosomal protein S12 mutations on cell-free protein synthesis. *European Journal of Biochemistry*, 271, 1127-1134.
- CLINICALTRIALS.GOV. 2020. Search Results for vaccine, Phase 1-4, U.S. [Online]. Available: <https://clinicaltrials.gov/ct2/results/map?term=vaccine&phase=0123&map=> [Accessed July 24, 2020].
- COHEN, B. J. & RICHMOND, J. E. 1982. Electron microscopy of hepatitis B core antigen synthesized in E. coli. *Nature*, 296, 677-678.
- COLE, S. D., BEABOUT, K., TURNER, K. B., SMITH, Z. K., FUNK, V. L., HARBAUGH, S. V., LIEM, A. T., ROTH, P. A., GEIER, B. A., EMANUEL, P. A., WALPER, S. A., CHÁVEZ, J. L. & LUX, M. W. 2019. Quantification of Interlaboratory Cell-Free Protein Synthesis Variability. *ACS Synthetic Biology*, 8, 2080-2091.
- CROWELL, L. E., LU, A. E., LOVE, K. R., STOCKDALE, A., TIMMICK, S. M., WU, D., WANG, Y., DOHERTY, W., BONNYMAN, A., VECCHIARELLO, N., GOODWINE, C., BRADBURY, L., BRADY, J. R., CLARK, J. J., COLANT, N. A., CVETKOVIC, A., DALVIE, N. C., LIU, D., LIU, Y.,

- MASCARENHAS, C. A., MATTHEWS, C. B., MOZDZIERZ, N. J., SHAH, K. A., WU, S.-L., HANCOCK, W. S., BRAATZ, R. D., CRAMER, S. M. & LOVE, J. C. 2018. On-demand manufacturing of clinical-quality biopharmaceuticals. *Nature Biotechnology*, 36, 988-995.
- DE GROOT, N. S. & VENTURA, S. 2006. Effect of temperature on protein quality in bacterial inclusion bodies. *FEBS Letters*, 580, 6471-6476.
- DES SOYE, B. J., GERBASI, V. R., THOMAS, P. M., KELLEHER, N. L. & JEWETT, M. C. 2019. A highly productive, one-pot cell-free protein synthesis platform based on genomically recoded *Escherichia coli*. *Cell chemical biology*, 26, 1743-1754. e9.
- DIAZ, D., VIDAL, X., SUNNA, A. & CARE, A. 2020. Bioengineering a light-responsive encapsulin nanoreactor: a potential tool for photodynamic therapy. *bioRxiv*, 2020.06.06.138305.
- DOERR, A., DE REUS, E., VAN NIES, P., VAN DER HAAR, M., WEI, K., KATTAN, J., WAHL, A. & DANELON, C. 2019. Modelling cell-free RNA and protein synthesis with minimal systems. *Physical biology*, 16, 025001.
- DOPP, J. L., JO, Y. R. & REUEL, N. F. 2019a. Methods to reduce variability in *E. Coli*-based cell-free protein expression. *Synthetic and Systems Biotechnology*, 4, 204-211.
- DOPP, J. L. & REUEL, N. F. 2018. Process optimization for scalable *E. coli* extract preparation for cell-free protein synthesis. *Biochemical Engineering Journal*, 138, 21-28.
- DOPP, J. L. & REUEL, N. F. 2019. One-Pot Cell-Free Extract for *in vitro* Expression of Disulfide Bonded Proteins. *bioRxiv*, 2019.12.19.883413.
- DOPP, J. L., TAMIEV, D. D. & REUEL, N. F. 2019b. Cell-free supplement mixtures: Elucidating the history and biochemical utility of additives used to support *in vitro* protein synthesis in *E. coli* extract. *Biotechnology Advances*, 37, 246-258.
- DORON, N. Bacterial Strains for Protein Production. The Wolfson Centre for Applied Structural Biology at the Hebrew University of Jerusalem.
- DOU, D., REVOL, R., OSTBYE, H., WANG, H. & DANIELS, R. 2018. Influenza A Virus Cell Entry, Replication, Virion Assembly and Movement. *Frontiers in Immunology*, 9.
- ENDO, Y., OTSUZUKI, S., ITO, K. & MIURA, K.-I. 1992. Production of an enzymatic active protein using a continuous flow cell-free translation system. *Journal of biotechnology*, 25, 221-230.
- ENGINZYME. 2020. *Our Technology* [Online]. Available: <http://enginzyne.com/technology/> [Accessed].
- EVROGEN. 2020. *Green fluorescent protein TurboGFP* [Online]. Available: [http://evrogen.com/products/TurboGFP/TurboGFP\\_Detailed\\_description.s.html](http://evrogen.com/products/TurboGFP/TurboGFP_Detailed_description.s.html) [Accessed].

- FAILMEZGER, J., RAUTER, M., NITSCHER, R., KRAML, M. & SIEMANN-HERZBERG, M. 2017. Cell-free protein synthesis from non-growing, stressed *Escherichia coli*. *Nature Scientific Reports*, 7, 1-10.
- FAN, X., HASHEM, A. M., CHEN, Z., LI, C., DOYLE, T., ZHANG, Y., YI, Y., FARNSWORTH, A., XU, K., LI, Z., HE, R., LI, X. & WANG, J. 2015. Targeting the HA2 subunit of influenza A virus hemagglutinin via CD40L provides universal protection against diverse subtypes. *Mucosal Immunology*, 8, 211-220.
- FAREWELL, A. & NEIDHARDT, F. C. 1998. Effect of Temperature on In Vivo Protein Synthetic Capacity in *Escherichia coli*. *Journal of Bacteriology*, 180, 4704-4710.
- FDA. 2020. *Vaccines Licensed for Use in the United States* [Online]. Available: <https://www.fda.gov/vaccines-blood-biologics/vaccines/vaccines-licensed-use-united-states> [Accessed July 24, 2020].
- FDA Decemeber 19, 2017. FDA Approves Novel Gene Therapy to Treat Patients with Rare Form of Inherited Vision Loss. FDA News Release.
- FOSHAG, D., HENRICH, E., HILLER, E., SCHÄFER, M., KERGER, C., BURGER-KENTISCHER, A., DIAZ-MORENO, I., GARCÍA-MAURIÑO, S. M., DÖTSCH, V., RUPP, S. & BERNHARD, F. 2018. The *E. coli* S30 lysate proteome: A prototype for cell-free protein production. *New Biotechnology*, 40, 245-260.
- FPBASE 2020a. Folding Reporter GFP.
- FPBASE. 2020b. *Superfolder GFP* [Online]. Available: <https://www.fpbases.org/protein/superfolder-gfp/> [Accessed].
- FRANCIS, J. N., BUNCE, C. J., HORLOCK, C., WATSON, J. M., WARRINGTON, S. J., GEORGES, B. & BROWN, C. B. 2015. A novel peptide-based pan-influenza A vaccine: a double blind, randomised clinical trial of immunogenicity and safety. *Vaccine*, 33, 396-402.
- FRANCIS, M. J. 2018. Recent Advances in Vaccine Technologies. *The Veterinary clinics of North America. Small animal practice*, 48, 231-241.
- GARCÍA-FRUITÓS, E. 2015. *Insoluble proteins*, Springer.
- GEORGI, V., GEORGI, L., BLECHERT, M., BERGMEISTER, M., ZWANZIG, M., WUSTENHAGEN, D. A., BIER, F. F., JUNG, E. & KUBICK, S. 2016. On-chip automation of cell-free protein synthesis: new opportunities due to a novel reaction mode. *Lab Chip*, 16, 269-281.
- GERDIL, C. 2003. The annual production cycle for influenza vaccine. *Vaccine*, 21, 1776–1779.
- GOERKE, A. R. & SWARTZ, J. R. 2008. Development of cell-free protein synthesis platforms for disulfide bonded proteins. *Biotechnology and Bioengineering*, 99, 351-367.

- GOLDRICK, S., HOLMES, W., BOND, N. J., LEWIS, G., KUIPER, M., TURNER, R. & FARID, S. S. 2017. Advanced multivariate data analysis to determine the root cause of trisulfide bond formation in a novel antibody-peptide fusion. *Biotechnology and bioengineering*, 114, 2222-2234.
- GOLDRICK, S., SANDNER, V., CHEEKS, M., TURNER, R., FARID, S. S., MCCREATH, G. & GLASSEY, J. 2020. Multivariate Data Analysis Methodology to Solve Data Challenges Related to Scale-Up Model Validation and Missing Data on a Micro-Bioreactor System. *Biotechnology Journal*, 15, 1800684.
- GONG, P. & MARTIN, C. T. 2006. Mechanism of instability in abortive cycling by T7 RNA polymerase. *Journal of Biological Chemistry*, 281, 23533-23544.
- GREGORIO, N. E., LEVINE, M. Z. & OZA, J. P. 2019. A user's guide to cell-free protein synthesis. *Methods and protocols*, 2, 24.
- HANSEN, M. M. K., ROSQUELLES, M. V., YELLESWARAPU, M., MAAS, R. J. M., VAN VUGT-JONKER, A. J., HEUS, H. A. & HUCK, W. T. S. 2016. Protein Synthesis in Coupled and Uncoupled Cell-Free Prokaryotic Gene Expression Systems. *ACS Synthetic Biology*, 5, 1433-1440.
- HAUSE, B. M., COLLIN, E. A., LIU, R., HUANG, B., SHENG, Z., LU, W., WANG, D., NELSON, E. A. & LI, F. 2014. Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *mBio*, 5, e00031-14.
- HOELSCHER, M. A., GARG, S., BANGARI, D. S., BELSER, J. A., LU, X., STEPHENSON, I., BRIGHT, R. A., KATZ, J. M., MITTAL, S. K. & SAMBHARA, S. 2006. Development of adenoviral-vector-based pandemic influenza vaccine against antigenically distinct human H5N1 strains in mice. *The Lancet*, 367, 475-481.
- HOLMES, K., SHEPHERD, D. A., ASHCROFT, A. E., WHELAN, M., ROWLAND, D. J. & STONEHOUSE, N. J. 2015. Assembly Pathway of Hepatitis B Core Virus-like Particles from Genetically Fused Dimers. *Journal of Biological Chemistry*, 290, 16238-16425.
- HONG, S. H., KWON, Y.-C., MARTIN, R. W., DES SOYE, B. J., DE PAZ, A. M., SWONGER, K. N., NTAI, I., KELLEHER, N. L. & JEWETT, M. C. 2015. Improving cell-free protein synthesis through genome engineering of *Escherichia coli* lacking release factor 1. *Chembiochem : a European journal of chemical biology*, 16, 844-853.
- HONG, S. H., KWON, Y. C. & JEWETT, M. C. 2014. Non-standard amino acid incorporation into proteins using *Escherichia coli* cell-free protein synthesis. *Frontiers in Chemistry*, 2.
- HOQUE, M., SHIMIZU, N., ISHIZU, K.-I., YAJIMA, H., ARISAKA, F., SUZUKI, K., WATANABE, H. & HANDA, H. 1999. Chimeric Virus-like Particle Formation of Adeno-Associated Virus. *Biochemical and Biophysical Research Communications*, 266, 371-376.

- HUB, U. F. T. H. M. 2016. Future Targeted Healthcare Manufacturing Hub. University College London.
- HUNT, J. P., YANG, S. O., WILDING, K. M. & BUNDY, B. C. 2017. The growing impact of lyophilized cell-free protein expression systems. *Bioengineered*, 8, 325-330.
- ICH 2009. ICH Harmonised Tripartite Guideline. Q8 (R2) *Current Step*, 4.
- ISKAKOVA, M. B., SZAFLARSKI, W., DREYFUS, M., REMME, J. & NIERHAUS, K. H. 2006. Troubleshooting coupled in vitro transcription-translation system derived from *Escherichia coli* cells: synthesis of high-yield fully active proteins. *Nucleic Acids Research*, 34.
- JEWETT, M. C., CALHOUN, K. A., VOLOSHIN, A., WUU, J. J. & SWARTZ, J. R. 2008. An integrated cell-free metabolic platform for protein production and synthetic biology. *Molecular systems biology*, 4.
- JEWETT, M. C., FRITZ, B. R., TIMMERMAN, L. E. & CHURCH, G. M. 2013. In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Molecular Systems Biology*, 9, 678.
- JEWETT, M. C., HODGMAN, C. E. & GAN, R. 2016a. *Methods for cell-free protein synthesis*. U.S. patent application.
- JEWETT, M. C., MARTIN, R. W., HONG, S. H., KWON, Y. C. & DES SOYE, B. J. 2016b. *Methods for improved in vitro protein synthesis with proteins containing non standard amino acids*. U.S. patent application.
- JEWETT, M. C. & SWARTZ, J. R. 2004. Mimicking the *Escherichia coli* Cytoplasmic Environment Activates Long-Lived and Efficient Cell-Free Protein Synthesis. *Biotechnology Bioengineering*, 86, 19-26.
- JIANG, L., ZHAO, J., LIAN, J. & XU, Z. 2018. Cell-free protein synthesis enabled rapid prototyping for metabolic engineering and synthetic biology. *Synthetic and Systems Biotechnology*, 3, 90-96.
- JIN, X. & HONG, S. H. 2018. Cell-free protein synthesis for producing 'difficult-to-express' proteins. *Biochemical Engineering Journal*, 138, 156-164.
- JUNGE, F., HABERSTOCK, S., ROOS, C., STEFER, S., PROVERBIO, D., DOTSCHE, V. & BERNHARD, F. 2011. Advances in cell-free protein synthesis for the functional and structural analysis of membrane proteins. *New Biotechnology*, 28, 262-271.
- KAI, L., DOTSCHE, V., KALDENHOFF, R. & BERNHARD, F. 2013. Artificial environments for the co-translational stabilization of cell-free expressed proteins. *PLoS One*, 8.
- KAZAKS, A., LU, I. N., FARINELLE, S., RAMIREZ, A., CRESCENTE, V., BLAHA, B., OGANAH, O., MUKHOPADHYAY, T., DE OBANOS, M. P., KRIMER, A., AKOPJANA, I., BOGANS, J., OSE, V., KIRSTEINA, A., KAZAKA, T., STONEHOUSE, N. J., ROWLANDS, D. J., MULLER, C. P., TARS, K. & ROSENBERG, W. M. 2017. Production and purification of chimeric HBc

- virus-like particles carrying influenza virus LAH domain as vaccine candidates. *BMC Biotechnology*, 17, 79.
- KIM, D. M. & CHOI, C. Y. 1996. A Semicontinuous Prokaryotic Coupled Transcription/Translation System Using a Dialysis Membrane. *Biotechnology Progress*, 12, 645-649.
- KIM, D. M. & SWARTZ, J. R. 1999. Prolonging Cell-Free Protein Synthesis with a Novel ATP Regeneration System. *Biotechnology and Bioengineering*, 66, 180-188.
- KIM, T.-W., KIM, D.-M. & CHOI, C.-Y. 2006a. Rapid production of milligram quantities of proteins in a batch cell-free protein synthesis system. *Journal of Biotechnology*, 124, 373-380.
- KIM, T. W., KEUM, J. W., OH, I. S., CHOI, C. Y., PARK, C. G. & KIM, D. M. 2006b. Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *Journal of Biotechnology*, 126, 554-561.
- KING, J. A., DUBIELZIG, R., GRIMM, D. & KLEINSCHMIDT, J. A. 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO journal*, 20, 3282-3291.
- KNAPP, K. G., GOERKE, A. R. & SWARTZ, J. R. 2006. Cell-Free Synthesis of Proteins That Require Disulfide Bonds Using Glucose as an Energy Source. *Biotechnology and Bioengineering*, 97, 901-908.
- 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. *Biotechnology Letters*, 34, 67-73.
- KRINSKY, N., KADURI, M., SHAINSKY-ROITMAN, J., GOLDFEDER, M., IVANAIR, E., BENHAR, I., SHOHAM, Y. & SCHROEDER, A. 2016. A Simplified and Rapid Method for Preparing a Cell-Free Bacterial Lysate for Protein Synthesis. *PLoS ONE*, 11.
- KRZYSZCZYK, P., ACEVEDO, A., DAVIDOFF, E. J., TIMMINS, L. M., MARRERO-BERRIOS, I., PATEL, M., WHITE, C., LOWE, C., SHERBA, J. J., HARTMANSHEEN, C., O'NEILL, K. M., BALTER, M. L., FRITZ, Z. R., ANDROULAKIS, I. P., SCHLOSS, R. S. & YARMUSH, M. L. 2018. The growing role of precision and personalized medicine for cancer treatment. *Technology*, 6, 79-100.
- KWON, Y. C. & JEWETT, M. C. 2015. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Nature Scientific Reports*, 5, 1-8.
- LE, D. T., RADUKIC, M. T. & MÜLLER, K. M. 2019. Adeno-associated virus capsid protein expression in *Escherichia coli* and chemically defined capsid assembly. *Scientific reports*, 9, 1-10.

- LEE, K.-Y., LEE, K.-H., PARK, J.-W. & KIM, D.-M. 2012. Flexible programming of cell-free protein synthesis using magnetic bead-immobilized plasmids. *PLoS one*, 7, e34429-e34429.
- LEGA, T., WEIHER, P., OBUCHOWSKI, M. & NIDZWORSKI, D. 2016. Presenting influenza A M2e Antigen on Recombinant Spores of *Bacillus subtilis*. *PLoS ONE*, 11.
- LEVINE, M. Z., SO, B., MULLIN, A. C., WATTS, K. R. & OZA, J. P. 2019. Redesigned upstream processing enables a 24-hour workflow from *E. coli* cells to cell-free protein synthesis. *bioRxiv*.
- LI, T., LI, C., QUAN, D. N., BENTLEY, W. E. & WANG, L.-X. 2018. Site-specific immobilization of endoglycosidases for streamlined chemoenzymatic glycan remodeling of antibodies. *Carbohydrate research*, 458, 77-84.
- LIU, D. V., ZAWADA, J. F. & SWARTZ, J. R. 2005. Streamlining Escherichia Coli S30 Extract Preparation for Economical Cell-Free Protein Synthesis. *Biotechnology Progress*, 21, 460-465.
- LU, I. N., KIRSTEINA, A., FARINELLE, S., WILLIEME, S., TARS, K., MULLER, C. P. & KAZAKS, A. 2018. Structure and applications of novel influenza HA tri-stalk protein for evaluation of HA stem-specific immunity. *PLOS ONE*, 13, e0204776.
- LU, Y., CHAN, W., KO, B. Y., VANLANG, C. C. & SWARTZ, J. R. 2015. Assessing sequence plasticity of a virus-like nanoparticle by evolution toward a versatile scaffold for vaccines and drug delivery. *PNAS*, 112, 12360-12365.
- LUCIGEN 2018. ClearColi BL21 (DE3) Electrocompetent Cells.
- LUDGATE, L., LIU, K., LUCKENBAUGH, L., STRECK, N., ENG, S., VOITENLEITNER, C., DELANEY, W. E. & HU, J. 2016. Cell-Free Hepatitis B Virus Capsid Assembly Dependent on the Core Protein C-Terminal Domain and Regulated by Phosphorylation. *Journal of Virology*, 90, 5830-5844.
- MANINI, I., DOMNICH, A., AMICIZIA, D., ROSSI, S., POZZI, T., GASPARINI, R., PANATTO, D. & MONTOMOLI, E. 2015. Flucelvax (Optaflu) for seasonal influenza. *Expert Review of Vaccines*, 14, 789-804.
- MARTINEZ-GIL, L. & MINGARRO, I. 2015. Viroporins, examples of the two-stage membrane protein folding model. *Viruses*, 7, 3462-3482.
- MAURER, A. C., PACOURET, S., CEPEDA DIAZ, A. K., BLAKE, J., ANDRES-MATEOS, E. & VANDENBERGHE, L. H. 2018. The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell reports*, 23, 1817-1830.
- MAURER, P., JENNINGS, G. T., WILLERS, J., ROHNER, F., LINDMAN, Y., ROUBICEK, K., RENNER, W. A., MULLER, P. & BACHMANN, M. F. 2005. A therapeutic vaccine for nicotine dependence: preclinical efficacy,

and phase 1 safety and immunogenicity. *European Journal of Immunology*, 35, 2013-2040.

- MELINEK, B., COLANT, N. A., STAMATIS, C., LENNON, C., FARID, S. S., POLIZZI, K., CARVER, M. & BRACEWELL, D. G. 2020. Toward a Roadmap for Cell-Free Synthesis in Bioprocessing. *BioProcess International*.
- MIGUEZ, A. M., MCNERNEY, M. P. & STYCZYNSKI, M. P. 2019. Metabolic Profiling of Escherichia coli-Based Cell-Free Expression Systems for Process Optimization. *Industrial & Engineering Chemistry Research*, 58, 22472-22482.
- MOZDZANOWSKA, K., FENG, J., EID, M., KRAGOL, G., CUDIC, M., OTVOS, L. & GERHARD, W. 2003. Induction of influenza type A virus-specific resistance by immunization of mice with a synthetic multiple antigenic peptide vaccine that contains ectodomains of matrix protein 2. *Vaccine*, 21, 2616-2626.
- NAGARAJ, V. H., GREENE, J. M., SENGUPTA, A. M. & SONTAG, E. D. 2017. Translation inhibition and resource balance in the TX-TL cell-free gene expression system. *Synthetic Biology*, 2.
- NASO, M. F., TOMKOWICZ, B., PERRY, W. L. & STROHL, W. R. 2017. Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs*, 31, 317-334.
- NEIRYNCK, S., DEROO, T., SAELENS, X., VANLANDSCHOOT, P., JOU, W. M. & FIERS, W. 1999. A universal influenza A vaccine based on the extracellular domain of the M2 protein. *Nature Medicine*, 5, 1157-1163.
- NIETO, K., WEGHOFER, M., SEHR, P., RITTER, M., SEDLMEIER, S., KARANAM, B., SEITZ, H., MÜLLER, M., KELLNER, M., HÖRER, M., MICHAELIS, U., RODEN, R. B. S., GISSMANN, L. & KLEINSCHMIDT, J. A. 2012. Development of AAVLP(HPV16/31L2) particles as broadly protective HPV vaccine candidate. *PloS one*, 7, e39741-e39741.
- NIRENBERG, M. W. & MATTHAEI, J. H. 1961. The Dependence of Cell-Free Protein Synthesis in E. coli Upon Naturally Occurring or Synthetic Polyribonucleotides. *Proceedings of the National Academy of Science of the United States of America*, 47, 1588-1602.
- OGONAH, O. W., POLIZZI, K. M. & BRACEWELL, D. G. 2017. Cell free protein synthesis: a viable option for stratified medicines manufacturing? *Current Opinion in Chemical Engineering*, 18, 77-83.
- OSTROWSKY, J., ARPEY, M., MOORE, K., OSTERHOLM, M., FRIEDE, M., GORDON, J., HIGGINS, D., MOLTO-LOPEZ, J., SEALS, J. & BRESEE, J. 2020. Tracking progress in universal influenza vaccine development. *Current Opinion in Virology*, 40, 28-36.
- OUSSOREN, C., ZUIDEMA, J., CROMMELIN, D. J. A. & STORM, G. 1997. Lymphatic Uptake and Biodistribution of Liposomes After Subcutaneous

- Inject. II. Influence of Liposomal Size, Lipid Composition, and Lipid Dose. *Biochimica et Biophysica Acta*, 1328, 262-272.
- OVERKAMP, W., BEILHARZ, K., DETERT OUDE WEME, R., SOLOPOVA, A., KARSENS, H., KOVACS, A., KOK, J., KUIPERS, O. P. & VEENING, J. W. 2013. Benchmarking various green fluorescent protein variants in *Bacillus subtilis*, *Streptococcus pneumoniae*, and *Lactococcus lactis* for live cell imaging. *Appl Environ Microbiol*, 79, 6481-90.
- OZAWA, K., JERGIC, S., CROWTHER, J. A., THOMPSON, P. R., WIJFFELS, G., OTTING, G. & DIXON, N. A. 2005. Cell-free Protein Synthesis in an Autoinduction System for NMR Studies of Protein-Protein Interactions. *Journal of Biomolecular NMR*, 32, 235-241.
- PALESE, P. & SHAW, M. L. 2007. Orthomyxoviridae: the viruses and their replication. *Fields Virology*. Philadelphia: Lippincott Williams & Wilkins.
- PANTHU, B., OHLMANN, T., PERRIER, J., SCHLATTNER, U., JALINOT, P., ELENA-HERRMANN, B. & RAUTUREAU, G. J. P. 2018. Cell-Free Protein Synthesis Enhancement from Real-Time NMR Metabolite Kinetics: Redirecting Energy Fluxes in Hybrid RRL Systems. *ACS Synthetic Biology*, 7, 218-226.
- PARDEE, K., GREEN, A. A., FERRANTE, T., CAMERON, D. E., DALEY KEYSER, A., YIN, P. & COLLINS, J. J. 2014. Paper-Based Synthetic Gene Networks. *Cell*, 159, 940-954.
- PARDEE, K., GREEN, A. A., TAKAHASHI, M. K., BRAFF, D., LAMBERT, G., WOOK LEE, J., FERRANTE, T., MA, D., DONGHIA, N., FAN, M., DARINGER, N. M., BOSCH, I., DUDLEY, D. M., O'CONNOR, D. H., GEHRKE, L. & COLLINS, J. J. 2016. Rapid, Low-Cost Detection of Zika Virus Using Programmable Biomolecular Components. *Cell*, 165, 1255-1266.
- PATEL, K. G. & SWARTZ, J. R. 2011. Surface functionalization of virus-like particles by direct conjugation using azide-alkyne click chemistry. *Bioconjugate chemistry*, 22, 376-387.
- PATTENDEN, L. K., MIDDLEBERG, A. P. J., NIEBERT, M. & LIPIN, D. 2005. Towards the preparative and large-scale precision manufacture of virus-like particles. *Trends in biotechnology*, 23, 523-529.
- PAULES, C. I., MARSTON, H. D., EISINGER, R. W., BALTIMORE, D. & FAUCI, A. S. 2017. The pathway to a universal influenza vaccine. *Immunity*, 47, 599-603.
- PEDELACQ, J. D., CABANTOUS, S., TRAN, T., TERWILLINGER, T. C. & WALDO, G. S. 2006. Engineering and characterization of a superfolder green fluorescent protein. *Nature Biotechnology*, 24, 79-88.
- PEDERSON, A., HELLBERG, K., ENBERG, J. & KARLSSON, B. G. 2011. Rational improvement of cell-free protein synthesis. *New Biotechnology*, 28, 218-224.

- PEREIRA, D. J., MCCARTY, D. M. & MUZYCZKA, N. 1997. The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *Journal of virology*, 71, 1079-1088.
- PESANT, S., NOT, F., PICHERAL, M., KANDELS-LEWIS, S., LE BESCOT, N., GORSKY, G., IUDICONE, D., KARSENTI, E., SPEICH, S., TROUBLÉ, R., DIMIER, C., SEARSON, S., ACINAS, S. G., BORK, P., BOSS, E., BOWLER, C., DE VARGAS, C., FOLLOWS, M., GORSKY, G., GRIMSLEY, N., HINGAMP, P., IUDICONE, D., JAILLON, O., KANDELS-LEWIS, S., KARP-BOSS, L., KARSENTI, E., KRZIC, U., NOT, F., OGATA, H., PESANT, S., RAES, J., REYNAUD, E. G., SARDET, C., SIERACKI, M., SPEICH, S., STEMMANN, L., SULLIVAN, M. B., SUNAGAWA, S., VELAYOUDON, D., WEISSENBACH, J., WINCKER, P. & TARA OCEANS CONSORTIUM, C. 2015. Open science resources for the discovery and analysis of Tara Oceans data. *Scientific Data*, 2, 150023.
- PEYRET, H., GEHLN, A., THUENEMANN, E. C., BLOND, D., TURABI, A. E., BEALES, L., CLARKE, D., GILBERT, R. J. C., FRY, E. E., STUART, D. I., HOLMES, K., STONEHOUSE, N. J., WHELAN, M., ROSENBERG, W., LOMONOSSOFF, G. P. & ROWLANDS, D. J. 2015. Tandem Fusion of Hepatitis B Core Antigen Allows Assembly of Virus-Like Particles in Bacteria and Plants with Enhanced Capacity to Accomodate Foreign Proteins. *PloS ONE*, 10, 1-20.
- PIELAK, R. M. & CHOU, J. J. 2011. Influenza M2 proton channels. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1808, 522-529.
- PREVENTION, C. F. D. C. A. 2019. Seasonal Influenza Vaccine Supply for the U.S. 2019-2020 Influenza Season.
- QUAST, R. B., KORTT, O., HENKEL, J., DONDAPATI, S. K., WÜSTENHAGEN, 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.
- RAMIREZ, A., MORRIS, S., MAUCOURANT, S., D'ASCANIO, I., CREZCENTE, V., LU, I., FARINELLE, S., MULLER, C. P., WHELAN, M. & ROSENBERG, W. 2018. A virus-like particle vaccine candidate for influenza A virus based on multiple conserved antigens presented on hepatitis B tandem core particles. *Vaccine*.
- RAMQVIST, T., ANDREASSON, K. & DALIANIS, T. 2007. Vaccination, immune and gene therapy based on virus-like particles against viral infections and cancer. *Expert Opinion on Biological Therapy*, 7, 997-1007.
- REZAEI, F., MIRSHAFIEY, A., SHAHMAHMOODI, S., SHOJA, Z., GHAVAMI, N. & MOKHTARI-AZAD, T. 2013. Influenza virus-like particle containing two different subtypes of hemagglutinin confers protection in mice against lethal challenge with A/PR8 (H1N1) and A/HK (H3N2) viruses. *Iranian Red Crescent Medical Journal*, 15, 75.

- RININGER, J., FENNELL, A., SCHOUKROUN-BARNES, L. & PETERSON, C. Capacity Analysis for Viral Vector Manufacturing: Is There Enough?
- ROBERTS, T. M., RUDOLF, F., MEYER, A., PELLAUX, R., WHITEHEAD, E., PANKE, S. & HELD, M. 2016. Identification and Characterisation of a pH-stable GFP. *Nature Scientific Reports*, 6.
- ROHOVIE, M. J., NAGASAWA, M. & SWARTZ, J. R. 2017. Virus-like particles: Next-generation nanoparticles for targeted therapeutic delivery. *Bioengineering & Translational Medicine*, 2, 43-57.
- ROLDÃO, A., MELLADO, M. C. M., CASTILHO, L. R., CARRONDO, M. J. T. & ALVES, P. M. 2010. Virus-like particles in vaccine development. *Expert Review of Vaccines*, 9, 1149-1176.
- ROSE, J. A., MAIZEL, J. V., JR., INMAN, J. K. & SHATKIN, A. J. 1971. Structural proteins of adenovirus-associated viruses. *Journal of virology*, 8, 766-770.
- ROSENBLUM, G. & COOPERMAN, B. S. 2014. Engine out of the chassis: cell-free protein synthesis and its uses. *FEBS letters*, 588, 261-268.
- SALEHI, A. S. M., SMITH, M. T., BENNETT, A. M., WILLIAMS, J. B., PITT, W. G. & BUNDY, B. C. 2016. Cell-free protein synthesis of a cytotoxic cancer therapeutic: Onconase production and a just-add-water cell-free system. *Biotechnology Journal*, 11, 274-281.
- SAMJI, T. 2009. Influenza A: Understanding the Viral Life Cycle. *Yale Journal of Biology and Medicine*, 82, 153-159.
- SAMPATH, V. 2018. Bacterial endotoxin-lipopolysaccharide; structure, function and its role in immunity in vertebrates and invertebrates. *Agriculture and Natural Resources*, 52, 115-120.
- SHELLEKENS, H., ALDOSARI, M., TALSMA, H. & MASTROBATTISTA, E. 2017. Making individualized drugs a reality. *Nature Biotechnology*, 35, 507-513.
- SCHINN, S. M., BROADBENT, A., BRADLEY, W. T. & BUNDY, B. C. 2016. Protein synthesis directly from PCR: progress and applications of cell-free protein synthesis with linear DNA. *New Biotechnology*, 33, 480-487.
- SCHOBORG, J. A., HERSHEWE, J. M., STARK, J. C., KIGHTLINGER, W., KATH, J. E., JAROENTOMEETCHAI, T., NATARAJAN, A., DELISA, M. P. & JEWETT, M. C. 2018. A cell-free platform for rapid synthesis and testing of active oligosaccharyltransferases. *Biotechnology and Bioengineering*, 115, 739-750.
- SCHUMACHER, J., BACIC, T., STARITZBICHLER, R., DANESCHDAR, M., KLAMP, T., ARNOLD, P., JÄGLE, S., TÜRECI, Ö., MARKL, J. & SAHIN, U. 2018. Enhanced stability of a chimeric hepatitis B core antigen virus-like-particle (HBcAg-VLP) by a C-terminal linker-hexahistidine-peptide. *Journal of Nanobiotechnology*, 16, 39.

- SCHUMANN, W. & FERREIRA, L. C. S. 2004. Production of recombinant proteins in *Escherichia coli*. *Genetics and Molecular Biology*, 27, 442-453.
- SEOW, Y. & WOOD, M. J. 2009. Biological gene delivery vehicles: Beyond viral vectors. *Molecular Therapy*, 17, 767-777.
- SHAO, W., PAUL, A., ABBASI, S., CHAHAL, P. S., MENA, J. A., MONTES, J., KAMEN, A. & PRAKASH, S. 2012. A novel polyethyleneimine-coated adeno-associated virus-like particle formulation for efficient siRNA delivery in breast cancer therapy: preparation and in vitro analysis. *International journal of nanomedicine*, 7, 1575-1586.
- SHENG, J., LEI, S., YUAN, L. & FENG, X. 2017. Cell-free protein synthesis of norovirus virus-like particles. *Royal Society of Chemistry Advances*, 7, 28837.
- SHIMIZU, Y. & UEDA, T. 2010. PURE Technology. *Methods in Molecular Biology*, 11-21.
- SHIN, J. & NOIREAUX, V. 2010. Research Efficient cell-free expression with the endogenous *E. Coli* RNA polymerase and sigma factor 70.
- SHIRBAGHAEI, Z. & BOLHASSANI, A. 2016. Different applications of virus-like particles in biology and medicine: Vaccination and delivery systems. *Biopolymers*, 105, 113-132.
- SITARAMAN, K., ESPOSITO, D., KLARMANN, G., LE GRICE, S. F., HARTLEY, J. L. & CHATTERJEE, D. K. 2004. A novel cell-free protein synthesis system. *Journal of Biotechnology*, 110, 257-263.
- SMITH, M. T., VARNER, C. T., BUSH, D. B. & BUNDY, B. C. 2012. The incorporation of the A2 protein to produce novel Q $\beta$  virus-like particles using cell-free protein synthesis. *Biotechnology Progress*, 28, 549-555.
- SMITH, W., ANDREWS, C. H. & LAIDLAW, P. P. 1933. A virus obtained from influenza patients. *Lancet*, 2, 66-68.
- SMUDA, J. W. & CARTER, B. J. 1991. Adeno-associated viruses having nonsense mutations in the capsid genes: Growth in mammalian cells containing an inducible amber suppressor. *Virology*, 184, 310-318.
- STAMATIS, C. F., S. 2020. Process economics evaluation of cell-free synthesis for the commercial manufacture of antibody drug conjugates. *Biotechnology Journal (Submitted)*.
- STARK, J. C., HUANG, A., NGUYEN, P. Q., DUBNER, R. S., HSU, K. J., FERRANTE, T. C., ANDERSON, M., KANAPSKYTE, A., MUCHA, Q., PACKETT, J. S., PATEL, P., PATEL, R. & QAO, D. 2018. BioBits Bright: A fluorescent synthetic biology education kit. *Science Advances*, 4, 1-10.
- STECH, M., QUAST, R. B., SACHSE, R., SCHULZE, C., WÜSTENHAGEN, D. A. & KUBICK, S. 2014. A continuous-exchange cell-free protein synthesis system based on extracts from cultured insect cells. *PloS one*, 9.

- STEINBACH, S., WISTUBA, A., BOCK, T. & KLEINSCHMIDT, J. A. 1997. Assembly of adeno-associated virus type 2 capsids in vitro. *Journal of General Virology*, 78, 1453-1462.
- STEPANENKO, O. V., STEPANENKO, O. V., KUZNETSOVA, I. M., VERKHUSHA, V. V. & TUROVEROV, K. K. 2014. Sensitivity of Superfolder GFP to Ionic Agents. *PLOS ONE*, 9, e110750.
- STEPHEN, S. L., BEALES, L., PEYRET, H., ROE, A., STONEHOUSE, N. J. & ROWLANDS, D. J. 2018. Recombinant Expression of Tandem-HBc Virus-Like Particles (VLPs). In: WEGE, C. & LOMONOSSOFF, G. P. (eds.) *Virus-Derived Nanoparticles for Advanced Technologies: Methods and Protocols*. New York, NY: Springer New York.
- STRYCHALSKI, E. A. & ROMANTSEVA, E. F. 2020. CELL-FREE (Comparable Engineered Living Lysates for Research Education and Entrepreneurship) Workshop Report.
- SUROSKY, R. T., URABE, M., GODWIN, S. G., MCQUISTON, S. A., KURTZMAN, G. J., OZAWA, K. & NATSOULIS, G. 1997. Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *Journal of virology*, 71, 7951-7959.
- TABOR, S. 1990. Expression using the T7 RNA polymerase/promoter system. *Current Protocols in Molecular Biology*, 11, 1-11.
- TEGERSTEDT, K., FRANZEN, A. V., ANDREASSON, K., JONEBERG, J., HEIDARI, S., RAMQVIST, T. & DALIANIS, T. 2005. Murine Polyomavirus Virus-like Particles (VLPs) as Vectors for Gene and Immune Therapy and Vaccines against Viral Infections and Cancer. *Anticancer Research*, 25, 2601-2608.
- THAVARAJAH, W., SILVERMAN, A. D., VEROSLOFF, M. S., KELLEY-LOUGHNANE, N., JEWETT, M. C. & LUCKS, J. B. 2020. Point-of-Use Detection of Environmental Fluoride via a Cell-Free Riboswitch-Based Biosensor. *ACS Synthetic Biology*, 9, 10-18.
- UNDERWOOD, K. A., SWARTZ, J. R. & PUGLISI, J. D. 2005. Quantitative Polysome Analysis Identifies Limitation in Bacterial Cell-Free Protein Synthesis. *Biotechnology and Bioengineering*, 91, 425-435.
- UNIVERCELLS. 2020. *NevoLine™ platform* [Online]. Available: <https://www.univercells.com/manufacturing-technologies/nevoline-platform/> [Accessed].
- VECCHIARELLO, N., TIMMICK, S. M., GOODWINE, C., CROWELL, L. E., LOVE, K. R., LOVE, J. C. & CRAMER, S. M. 2019. A combined screening and in silico strategy for the rapid design of integrated downstream processes for process and product-related impurity removal. *Biotechnology and bioengineering*, 116, 2178-2190.

- VOLOSHIN, A. M. & SWARTZ, J. R. 2005. Efficient and scalable method for scaling up cell free protein synthesis in batch mode. *Biotechnology and bioengineering*, 91, 516-521.
- WALKER, A., SKAMEL, C., VORREITER, J. & NASSAL, M. 2008. Internal core protein cleavage leaves the hepatitis B virus capsid intact and enhances its capacity for surface display of heterologous whole chain proteins. *Journal of Biological Chemistry*, 283, 33508-33515.
- WANG, D., TAI, P. W. L. & GAO, G. 2019a. Adeno-associated virus vector as a platform for gene therapy delivery. *Nature Reviews Drug Discovery*, 18, 358-378.
- WANG, P. H., FUJISHIMA, K., BERHANU, S., KURUMA, Y., JIA, T. Z., KHUSNUTDINOVA, A. N., YAKUNIN, A. F. & MCGLYNN, S. E. 2019b. A Single Polyphosphate Kinase-Based NTP Regeneration System Driving Cell-Free Protein Synthesis. *ChemRxiv*, Preprint.
- WARD, P., ELIAS, P. & LINDEN, R. M. 2003. Rescue of the adeno-associated virus genome from a plasmid vector: evidence for rescue by replication. *Journal of virology*, 77, 11480-11490.
- WHO 2013. Global Vaccine Action Plan 2011 - 2020.
- WILDING, K. M., HUNT, J. P., WILKERSON, J. W., FUNK, P. J., SWENSEN, R. L., CARVER, W. C., CHRISTIAN, M. L. & BUNDY, B. C. 2019. Endotoxin-Free E. coli-Based Cell-Free Protein Synthesis: Pre-Expression Endotoxin Removal Approaches for on-Demand Cancer Therapeutic Production. *Biotechnology Journal*, 14, 1-6.
- WRIGHT, J. F. 2014. Product-Related Impurities in Clinical-Grade Recombinant AAV Vectors: Characterization and Risk Assessment. *Biomedicines*, 2, 80-97.
- WU, P., XIAO, W., CONLON, T., HUGHES, J., AGBANDJE-MCKENNA, M., FERKOL, T., FLOTTE, T. & MUZYCZKA, N. 2000. Mutational Analysis of the Adeno-Associated Virus Type 2 (AAV2) Capsid Gene and Construction of AAV2 Vectors with Altered Tropism. *Journal of Virology*, 74, 8635-8647.
- WU, P. S. C., OZAWA, K., LIM, S. P., VASUDEVAN, S. G., DIXON, N. E. & OTTING, G. 2007. Cell-Free Transcription/Translation from PCR-Amplified DNA for High-Throughput NMR Studies. *Angewandte Chemie*, 46, 3356-3358.
- XIE, H., WAN, X.-F., YE, Z., PLANT, E. P., ZHAO, Y., XU, Y., LI, X., FINCH, C., ZHAO, N. & KAWANO, T. 2015. H3N2 mismatch of 2014–15 northern hemisphere influenza vaccines and head-to-head comparison between human and ferret antisera derived antigenic maps. *Scientific reports*, 5, 1-10.
- XU, Z., CHEN, H., YIN, X., XU, N. & CEN, P. 2005. High-level expression of soluble human  $\beta$ -defensin-2 fused with green fluorescent protein in

- Escherichia coli cell-free system. *Applied Biochemistry and Biotechnology*, 127, 53-62.
- YIN, G., GARCES, E. D., YANG, J., ZHANG, J., TRAN, C., STEINER, A. R., ROOS, C., BAJAD, S., HUDAK, S., PENTA, K., ZAWADA, J., POLLITT, S. & MURRAY, C. J. 2012. Aglycosylated antibodies and antibody gradients produced in a scalable in vitro transcription-translation system. *MAbs*, 4, 217-225.
- YOKOGAWA, T., KITAMURA, Y., NAKAMURA, D., OHNO, S. & NISHIKAWA, K. 2009. Optimization of the hybridization-based method for purification of thermostable tRNAs in the presence of tetraalkylammonium salts. *Nucleic Acids Research*, 38, e89-e89.
- ZAWADA, J. & SWARTZ, J. 2006. Effects of Growth Rate on Cell Extract Performance in Cell-Free Protein Synthesis. *Biotechnology and Bioengineering*, 94, 618-624.
- ZAWADA, J. F., YIN, G., STEINER, A. R., YANG, J., NARESH, A., ROY, S. M., GOLD, D. S., HEINSOHN, H. G. & MURRAY, C. J. 2011. Microscale to manufacturing scale-up of cell-free cytokine production—a new approach for shortening protein production development timelines. *Biotechnology and Bioengineering*, 108, 1570-1578.
- ZDANOWICZ, M. & CHROBOCZEK, J. 2016. Virus-like particles as drug delivery vectors. *Acta Biochimica Polonica*, 63, 469-473.
- ZELTINS, A. 2013. Construction and Characterization of Virus-Like Particles. *Molecular Biotechnology*, 53, 92-107.
- ZHANG, H.-G., XIE, J., DMITRIEV, I., KASHENTSEVA, E., CURIEL, D. T., HSU, H.-C. & MOUNTZ, J. D. 2002. Addition of six-His-tagged peptide to the C terminus of adeno-associated virus VP3 does not affect viral tropism or production. *Journal of virology*, 76, 12023-12031.
- ZIMMERMAN, E. S., HEIBECK, T. H., GILL, A., LI, X., MURRAY, C. J., MADLANSACAY, M. R., TRAN, C., UTER, N. T., YIN, G., RIVERS, P. J., YAM, A. Y., WANG, W. D., STEINER, A. R., BAJAD, S. U., PENTA, K., YANG, W., HALLAM, T. J., THANOS, C. D. & SATO, A. K. 2014. Production of site-specific antibody–drug conjugates using optimized non-natural amino acids in a cell-free expression system. *Bioconjugate Chemistry*, 25, 351-361.
- ZLOTNICK, A., PALMER, I., KAUFMAN, J. D., STAHL, S. J., STEVEN, A. C. & WINGFIELD, P. T. 1999. Separation and crystallization of T = 3 and T = 4 icosahedral complexes of the hepatitis B virus core protein. *Acta Crystallographica Section D*, 55, 717-720.
- ZUBAY, G. 1973. In vitro synthesis of protein in microbial systems. *Annual review of genetics*, 7, 267.



## 9. Appendix A: Protein Sequences and Plasmid Backbones

### 9.1 Hepatitis B Core Antigen Sequence

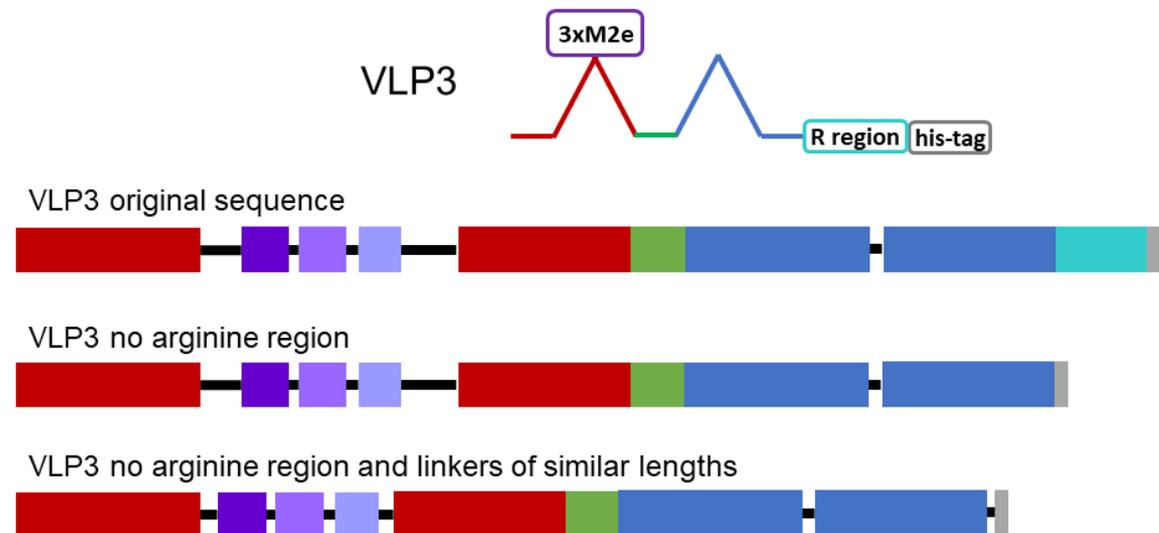
5'-MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDNASALYREALESPEHCSPHHT  
ALRQAILCWGELMTLATWVGGNLEDPISRDLVVS YVNTNMGLKFRQLLWFHISC  
LTFGRETVIEYLV SFGVWIRTPPAYRPPNAPILSTLPETT VVRRRGRSPRRRTPS  
PRRRRSQSPRRRRS QSRESQC-3'

### 9.2 Tandem-Core K1K1 VLP Sequence

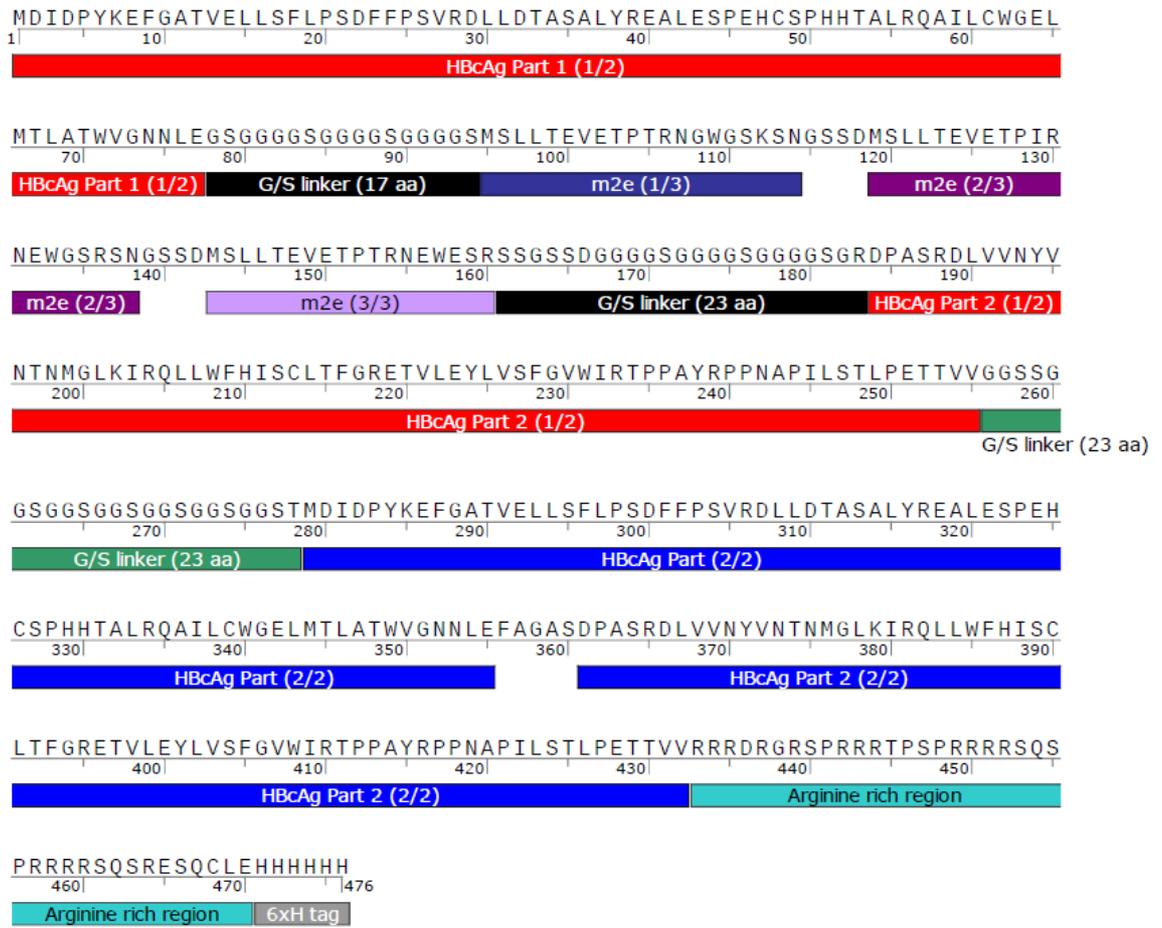
The lysine found in the MIRs are shown in bold:

5'-MDIDPYKEFGATVELLSFLPSDFFPSVRDLLDTASALYREALESPEHCSPHHT  
ALRQAILCWGELMTLATWVGN NLEGGSGSGSGGGG**K**GGGSGSSGRDPASRDLV  
VNYVNTNMGLKIRQLLWFHISCLTFGRETVLE YLV SFGVWIRTPPAYRPPNAPIL  
STLPETT VVGGSSGGSGGGSGGGSGGGSGGGSTMDIDPYKEFGATVELLSFLPS  
DFFPSVRDLLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGN  
NLEFGSGSGGG**K**GGGSGSASDPASRDLVNYVNTNMGLKIRQLLWFHISCLTF  
GRETVLE YLV SFGVWIRTPPAYRPPNAPILSTLPETT VVRRRDRGRSPRRRTPS  
PRRRRSQSPRRRRS QSRESQC-3'

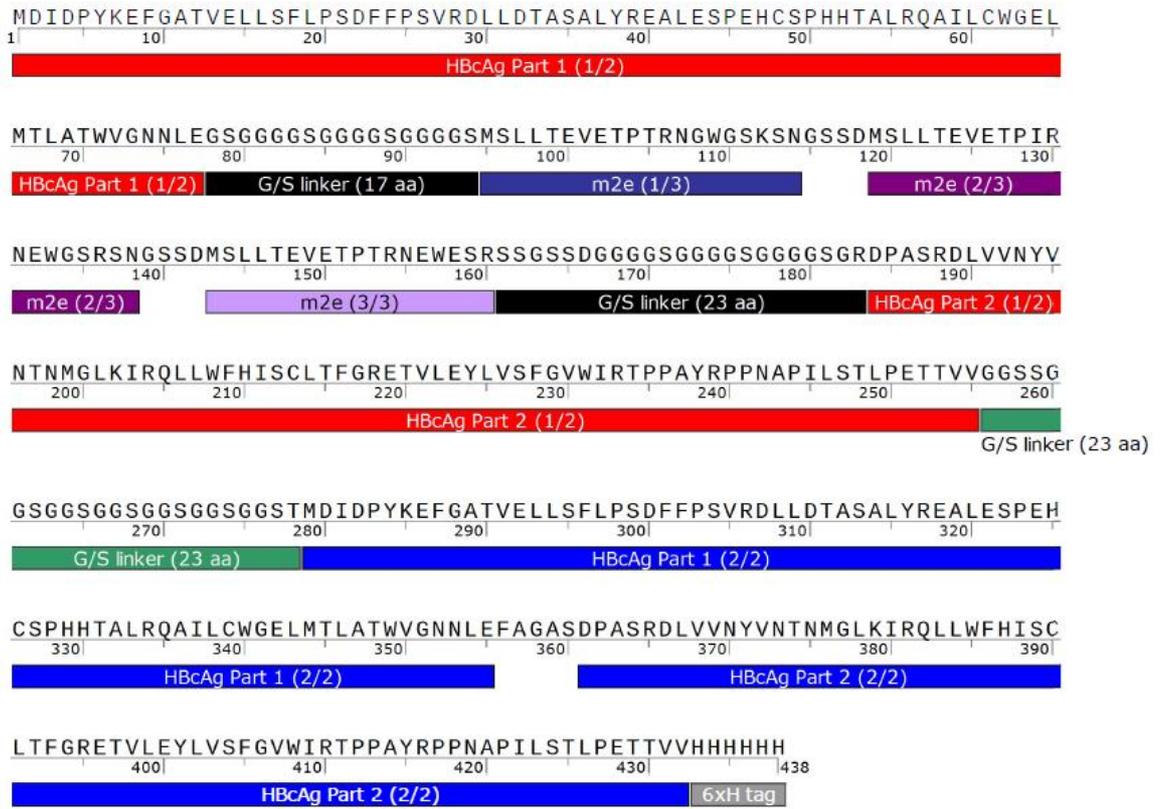
### 9.3 Tandem-Core IAV Vaccine Candidate Sequence and Annotated Modified Constructs



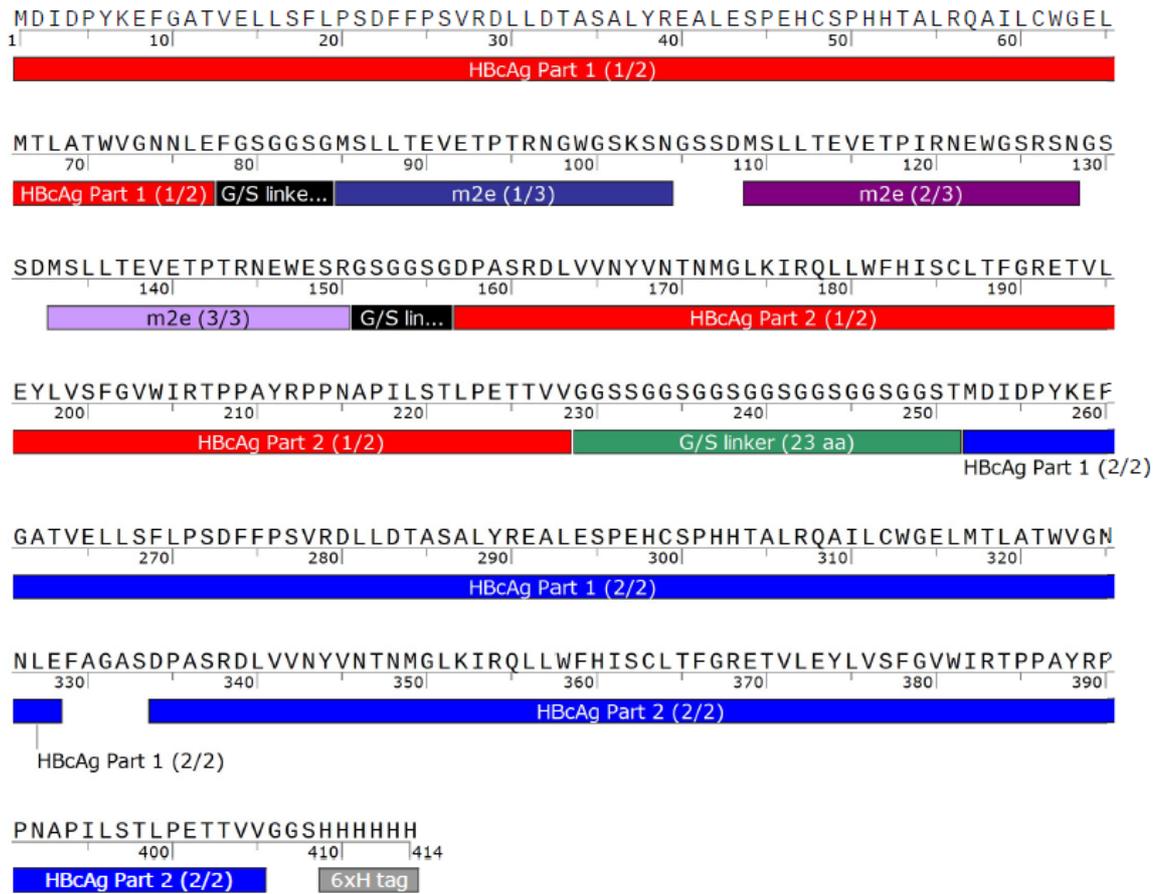
VLP 3 Original Protein Sequence:



VLP 3 No Arginine Region Protein Sequence:



VLP 3 No Arginine Region and Linkers of Simliar Lengths Protein Sequence:





VLP1 original sequence



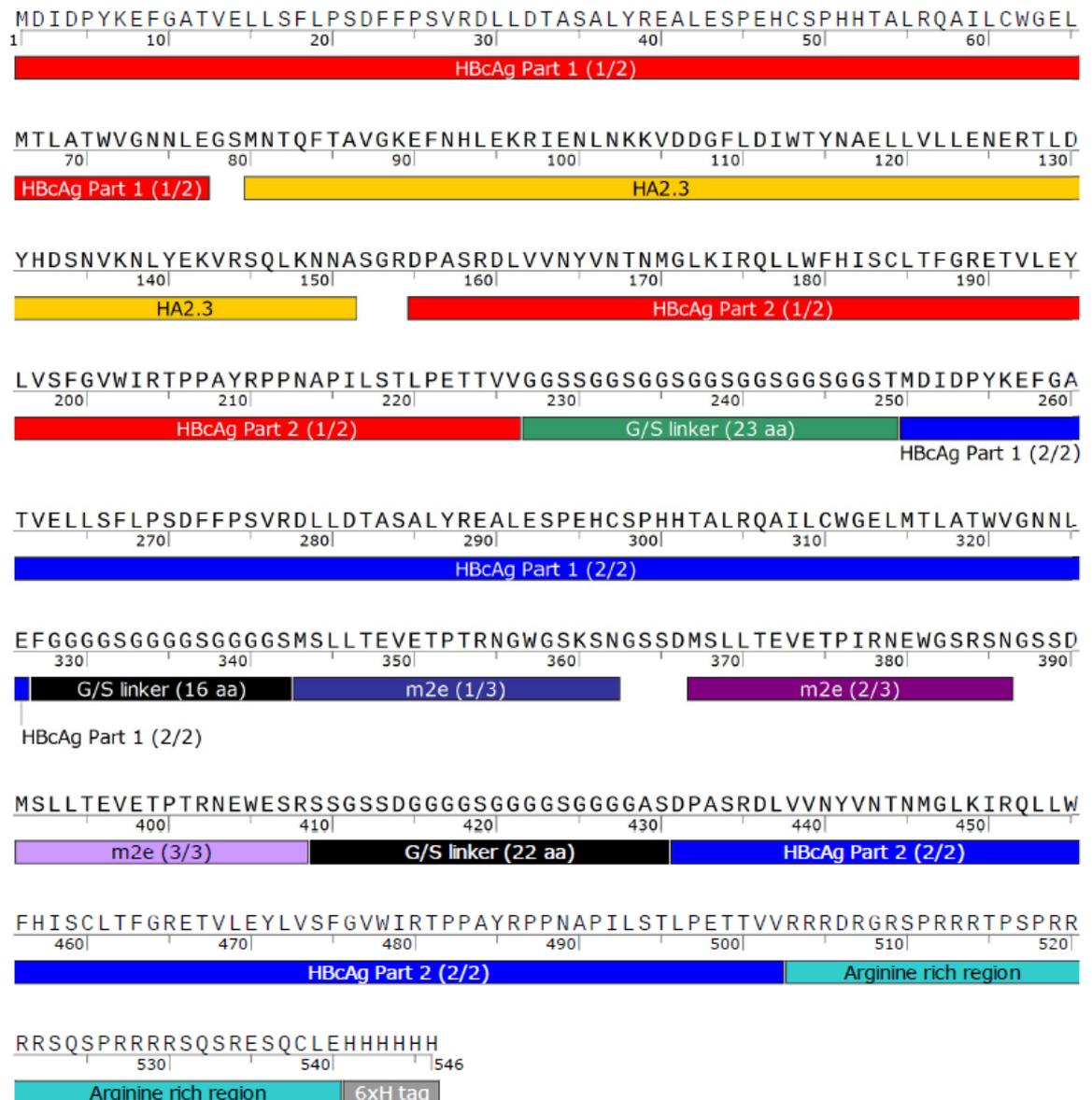
VLP1 no arginine region



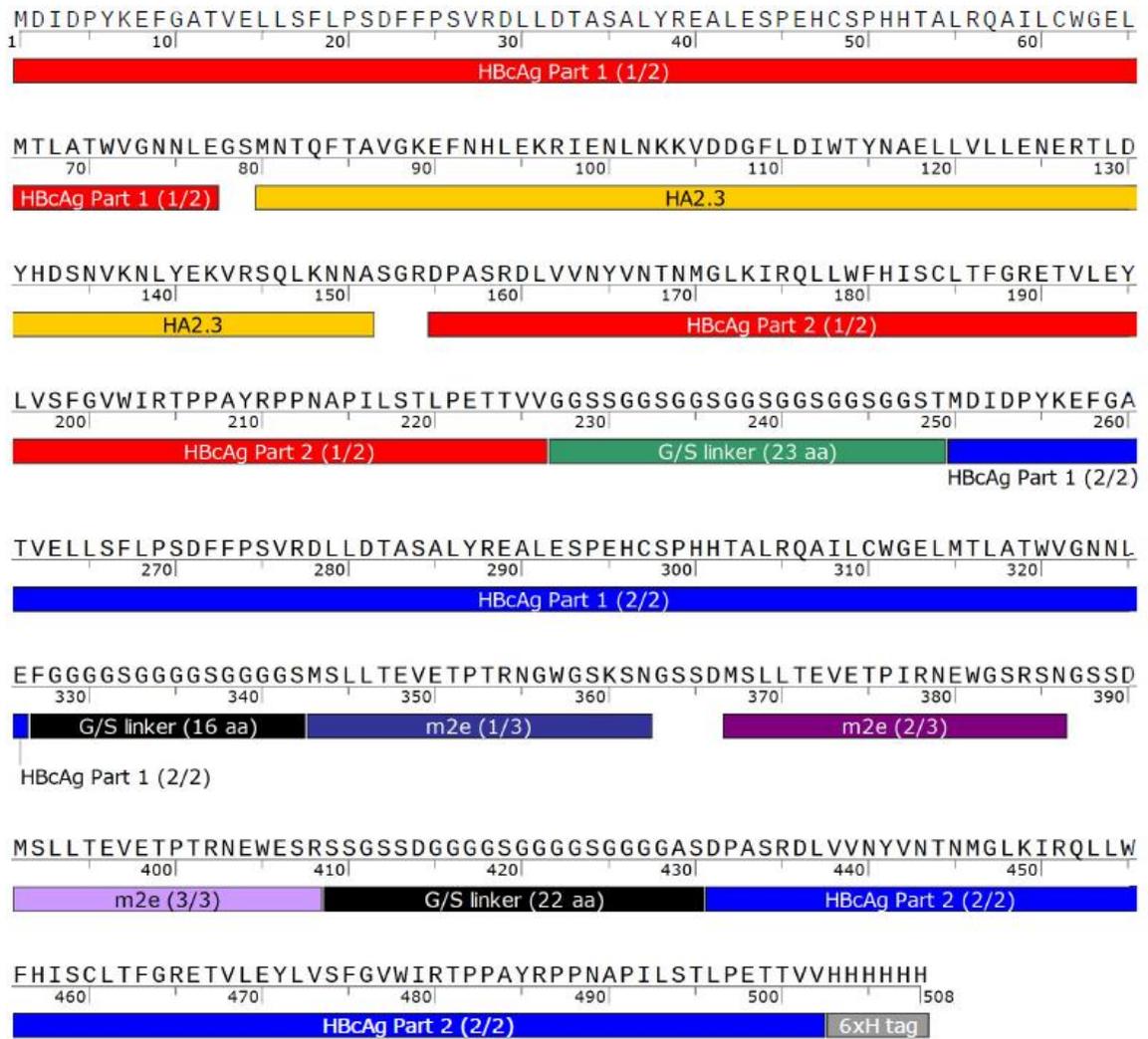
VLP1 no arginine region and linkers of similar lengths



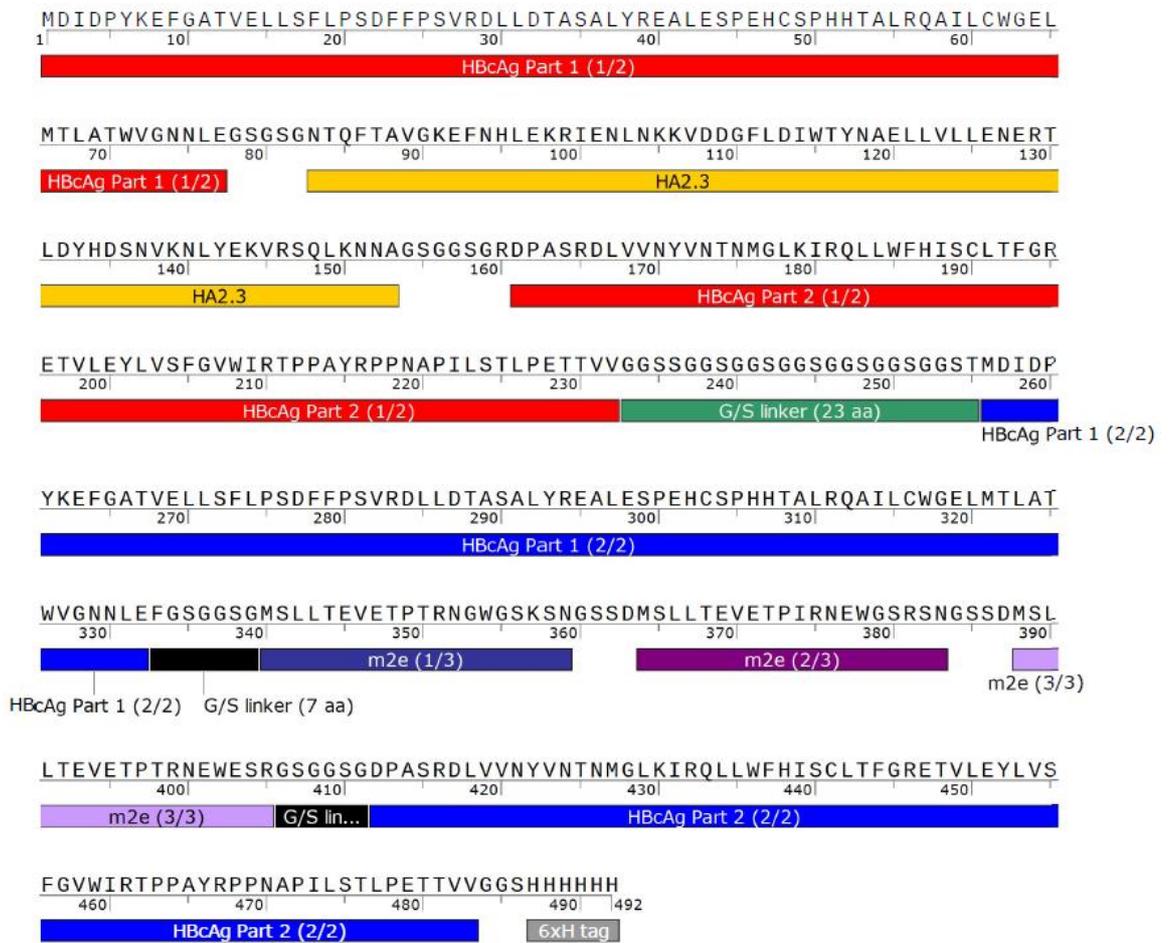
VLP 1 Original Protein Sequence:



VLP 1 No Arginine Region Protein Sequence:



VLP 1 No Arginine Region and Linkers of Similar Lengths Protein Sequence:



#### 9.4 Adeno-Associated Capsid Protein Sequences

6x histidine tags are showed in parentheses.

##### VP1

5'-M(HHHHHH)LVPRGSAADGYLPDWLEDTLSEGIRQWWKLPKPPPPKPAER  
HKDDSRGLVLPGYKYLGPFNGLDKGEPVNEADAAALEHDKAYDRQLDSDGNP  
YLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGK  
KRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDPQPLGQPPAA  
PSGLGTNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTR  
TVALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSPRDWQ  
RLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPY  
VLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLR  
TGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTNTPSGTTTQS  
RLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHL  
NGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEE  
IRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRDVYLQGP  
IWAKIPHTDGHFHPSPMLGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQ  
YSTGQVSVEIEWELQKENS KRWNPEIQYTSN YNKS VNVDFTVDTNGVYSEPRP  
IGTRYLTRNL-3'

##### VP2

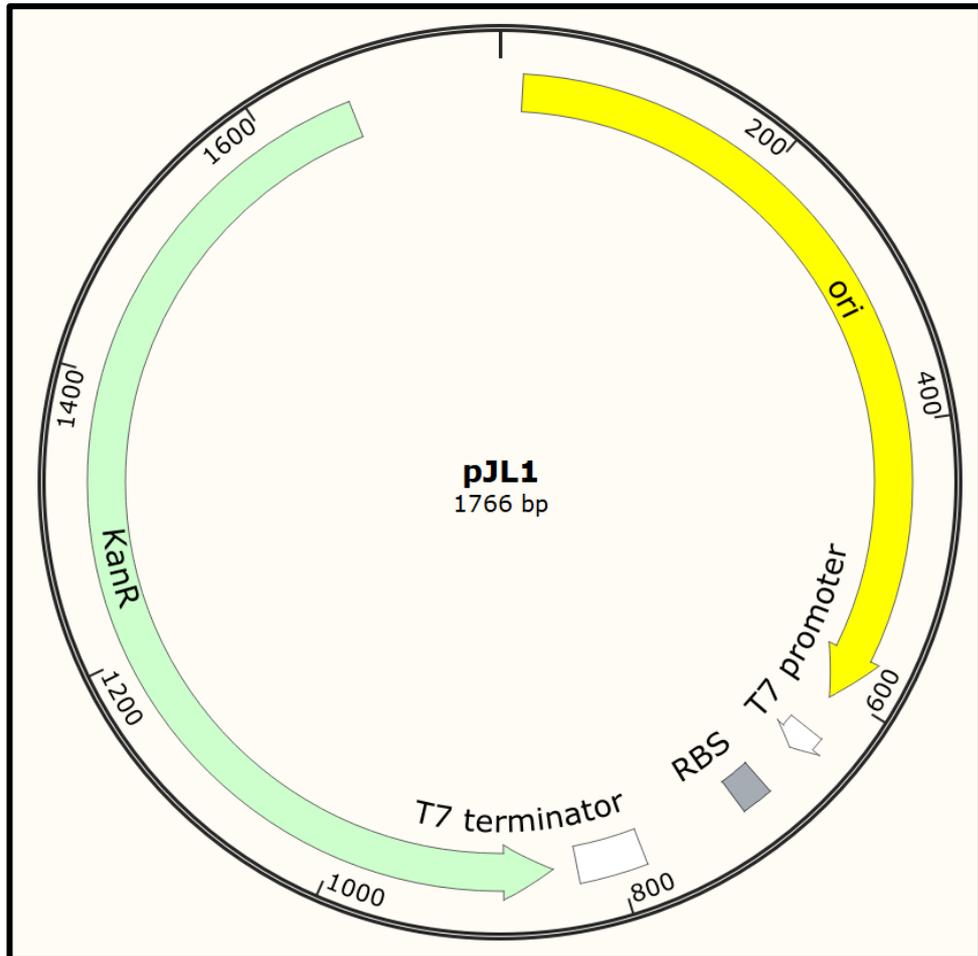
5'-M(HHHHHH)LVPRGSAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFG  
QTGDADSVDPDPQPLGQPPAAPSGGLGTNTMATGSGAPMADNNEGADGVGNSS  
GNWHCDSTWMGDRVITTSTRTVALPTYNNHLYKQISSQSGASNDNHYFGYST  
PWGYFDFNRFHCHFSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGT  
TTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNG  
SQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNP  
LIDQYLYLSRTNTPSGTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRV  
SKTSADNNNSEYSWTGATKYHLNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIF  
GKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADV  
NTQGVLPGMVWQDRDVYLQGP IWAKIPHTDGHFHPSPMLGGFGLKHPPPQILIK  
NTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSN  
YNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL-3'

##### VP3

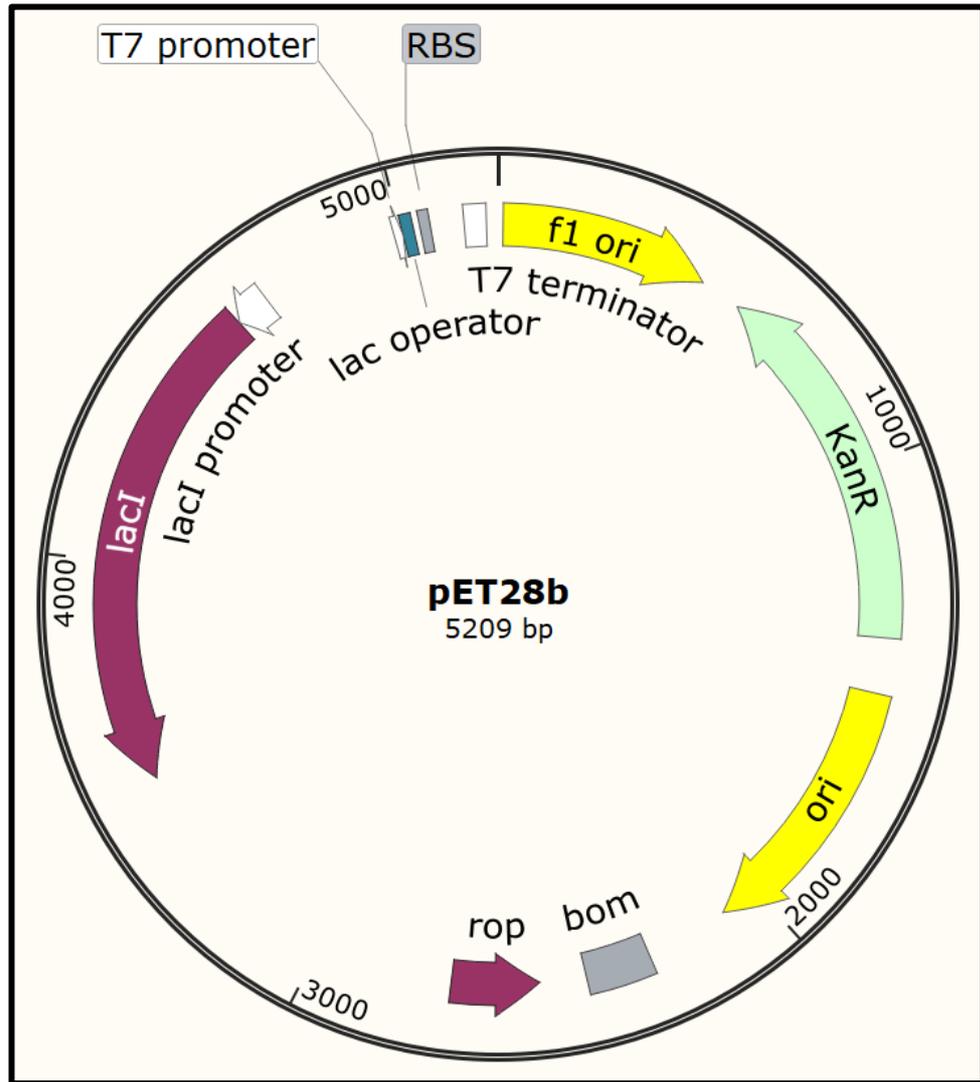
5'-M(HHHHHH)LVPRGSATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGD  
RVITTSTRTVALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCH  
FSPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFT  
DSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEY  
FPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTNTP  
SGTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWT  
GATKYHLNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEK  
VMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQD  
RDVYLQGP IWAKIPHTDGHFHPSPMLGGFGLKHPPPQILIKNTPVPANPSTTFSA  
AKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSN YNKS VNVDFTVDTN  
GVYSEPRPIGTRYLTRNL-3'

### 9.5 Plasmid Backbones

pJL1 backbone (optimised for CFPS)



pET28b backbone (*in vivo* expression plasmid)



## 10. Appendix B: Using CFPS to Produce “Difficult-to-Express” Products

### 10.1 CFPS for Expression of a Cold Environment Lipase

CFPS was used to express a 35 kDa lipase isolated from the Tara Oceans Database that is believed to have originated from cold-adapted bacteria found in the Antarctic ocean environment (Pesant et al., 2015). The lipase has been difficult to express under normal protein production conditions because its exact origins are unknown. Researchers in the UCL Biochemical Engineering department are modifying traditional expression platforms and exploring a variety of ways to produce this protein including increasing the salt concentration and decreasing the expression temperature based on similar work with other cold-adapted enzymes (Bjerga et al., 2016).

The CFPS system was used to test two expression temperatures, 10°C and 15°C, and to observe the differences between systems with high concentrations of IPTG and T7 RNA polymerase (IPTG-induced BL21-Star™ (DE3) cell extract) and low concentrations of IPTG and T7 RNA polymerase (non-induced BL21-Star™ cell extract). The reaction and the analysis were completed in less than a single working day. It was determined that more soluble protein was expressed in the non-induced BL21-Star™ cell extracts than the induced BL21-Star™ cell extracts and that this was true at both temperatures. This suggests that high concentrations of IPTG may be inhibiting production and that its concentration may need to be reduced *in vivo* for better expression. However, the control plasmid also shows expression at a similar molecular weight, suggesting that this is a host cell protein being expressed rather than the lipase.

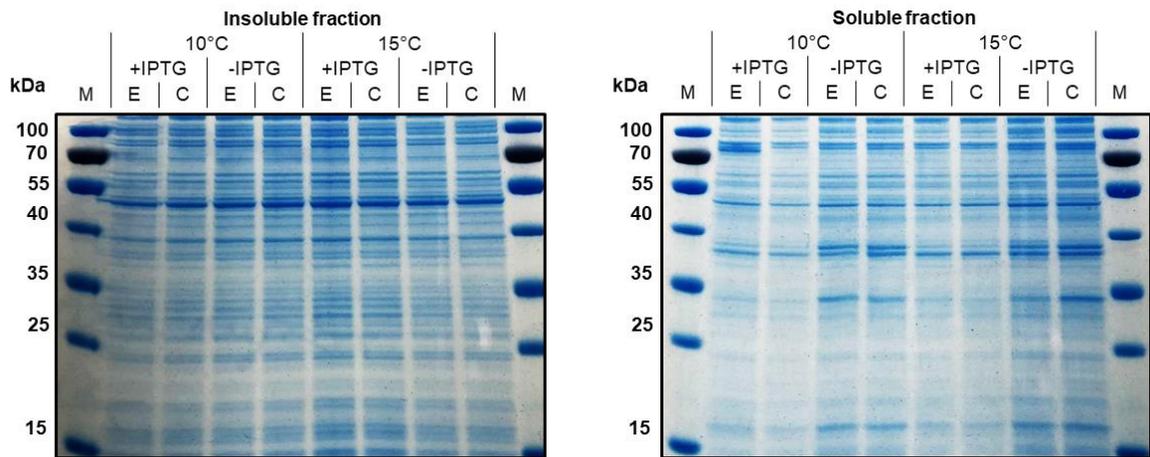


Figure 10.1 SDS-PAGE of CFPS Reactions for Cold Environment Lipase

Two different plasmids were tested in CFPS reactions, one for the 35 kDa cold environment lipase (E) and one for a “control” – the pET28a(+) plasmid without the lipase (C). They were tested in both BL21 Star™ induced (+IPTG) and non-induced (-IPTG) extracts at 10°C and 15°C.

### 10.2 CFPS of IgG-Specific Endoglycosidase

CFPS was used to express a 98 kDa endoglycosidase from *Streptococcus pyogenes* serotype M49 referred to as Endo-S2. Endo-S2 has been used previously for de-glycosylation of heterogeneous N-glycans to allow for more homogenous glycosylation patterns in IgG antibody products (Li et al., 2018). Three extracts were used: BL21-Star™ (DE3) that was induced with IPTG, non-induced BL21-Star™ (DE3), and IPTG-induced SHuffle T7. To promote disulphide bond formation, DsbC was added to a reaction using the IPTG-induced BL21-Star™ (DE3) extract that had been pre-treated with iodoacetamide. The protein was expressed in all reactions except the one containing DsbC. The lack of expression could be a result of incompatibilities between the buffers used for plasmid preparation and the buffer used to store the DsbC. Regardless, expression of Endo-S2 has now been demonstrated in a CFPS system.

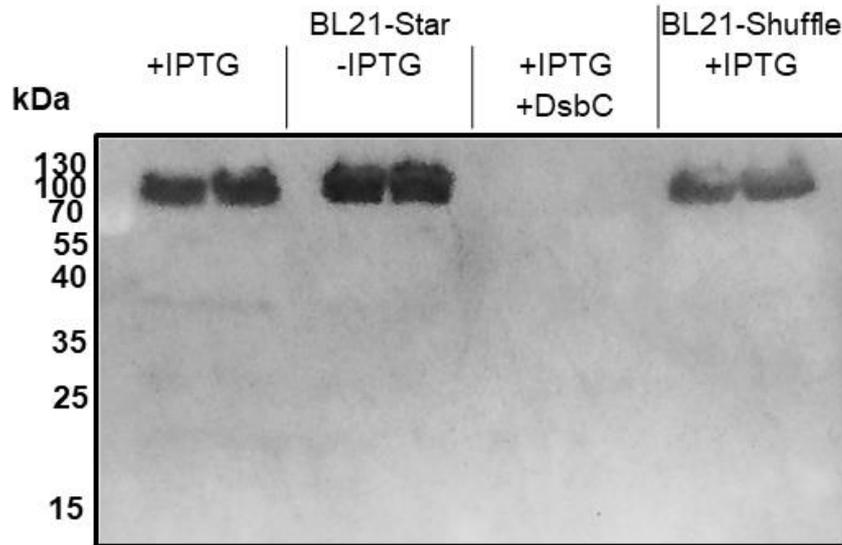


Figure 10.2 Western blot of ENDO-S2 CFPS Reactions

A plasmid expressing Endo-S2 with Strep-tag (98 kDa) was tested with BL21-Star<sup>TM</sup> extract that was either IPTG-induced (+IPTG) or non-induced (-IPTG) as well as an IPTG-induced SHuffle T7 extract. A reaction with DsbC was also performed.

### 10.3 FujiFilm Diosynth Biotechnologies User Feasibility Study

As part of a User Feasibility Study (UFS) for the FTHM Hub, the CFPS system developed in this doctoral project was transferred to FujiFilm Diosynth Biotechnologies UK Ltd. (FDB). The materials provided by UCL researchers for this UFS included: rTurbo GFP for standard curve preparation, noninduced BL21-Star<sup>TM</sup> cell extract, complex concentrated reaction mix, 50 mM amino acid solution (no methionine), 75 mM methionine solution, nuclease free water, and an aliquot of the pJL1 plasmid containing sfGFP. The reactions were performed in deep well flower plates at a volume of 1 mL. Samples were taken at 4 hours and 24 hours. The system was successful in synthesising sfGFP at levels previously observed at UCL (Figure 10.3).

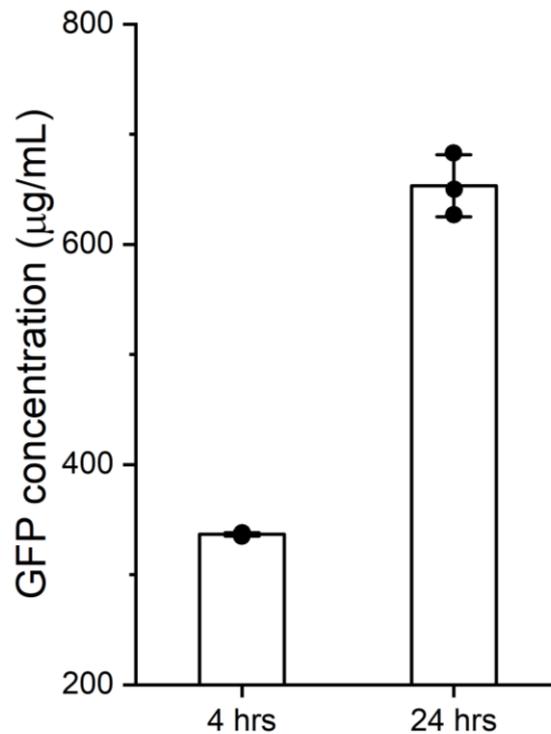


Figure 10.3 CFPS Reactions for FDB UFS

A noninduced BL21 Star™ (DE3) extract was combined with the concentrated reaction mixture based on the protocol by Kwon and Jewett (2015), additional T7 RNA polymerase, and the pJL1 plasmid expressing sfGFP. Samples were taken at 4 hours and 24 hours. Error bars represent plus or minus one standard deviation for n = 3 biological replicates.

Next, plasmids for difficult-to-express products prepared by researchers at FDB were tested in the reaction. No expression was detected (data not shown). There were three possible reasons that the difficult-to-express products were not synthesised in the CFPS reaction: 1) they were expressed but at a very low level because they are difficult-to-express, 2) the plasmid backbone used by FDB (not disclosed) is not compatible with the CFPS system, 3) the plasmid preparation methods used by FDB is not compatible with the CFPS system.

A researcher at FDB prepared a plasmid expressing sfGFP in the plasmid backbone used by FDB, tested it in the CFPS system, and could not detect any expression. This would indicate that the issue was most likely with the plasmid backbone or the plasmid preparation, rather than the product itself.

The sfGFP plasmid in the FDB backbone was transferred to UCL. It was transformed into JM109 cells and prepared using the Qiaprep Spin Miniprep Kit. When that plasmid was used in the CFPS reaction, sfGFP was expressed (Figure 10.4). This would indicate that the plasmid preparation technique used by FDB is not compatible with the CFPS system. The researchers at FDB prepared their plasmid onsite at FDB using Qiaprep Spin Miniprep Kit and re-eluted the previously prepared plasmid in a Qiaprep Spin Miniprep Kit Qiaspin column. When these plasmids were used in CFPS reactions, a green colour was visible. The exact reason for this incompatibility remains to be elucidated but this UFS has made evident the importance of the plasmid production process, an as of yet little studied area. It is possible that minor modification to the FDB plasmid preparation process, like a PCR clean up kit or a different elution buffer, could mitigate these incompatibilities and result in expression of the products designed by FDB.

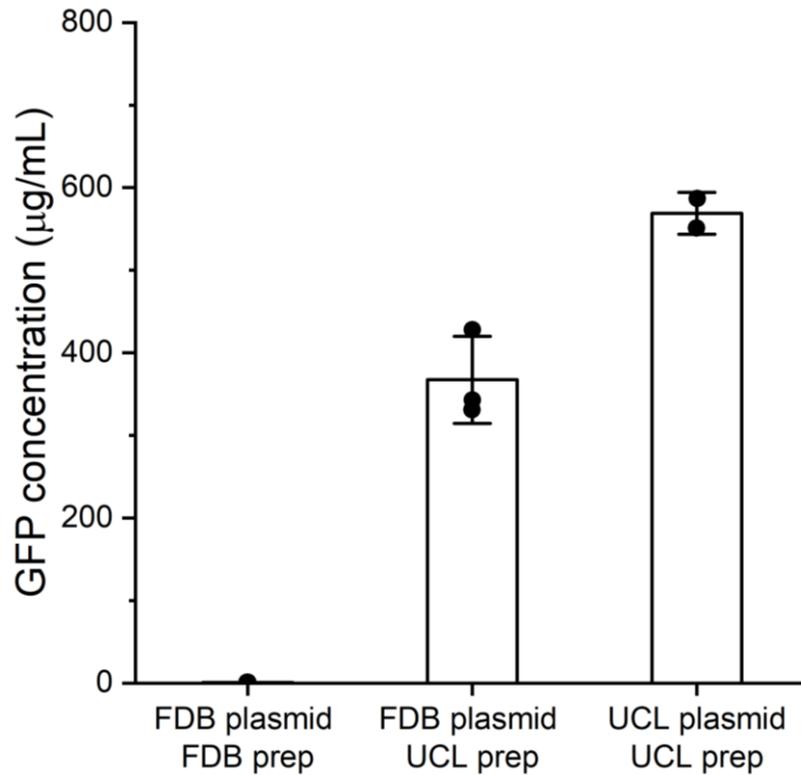


Figure 10.4 Comparing Plasmids Prepared at FDB and UCL

A noninduced BL21 Star™ (DE3) extract was combined with the concentrated reaction mixture based on the protocol by Kwon and Jewett (2015), additional T7 RNA polymerase, and either the plasmid expressing sfGFP in the backbone designed by FDB (FDB plasmid) or the pJL1 plasmid expressing sfGFP (UCL plasmid). The plasmids were either prepared at FDB (FDB prep) or at UCL (UCL prep). Samples were taken at 4 hours. Error bars represent plus or minus one standard deviation for  $n = 2$  technical replicates for all samples except the one with the FDB plasmid prepared at UCL which has  $n = 4$  technical replicates from 2 biological replicates.