



**This electronic thesis or dissertation has been
downloaded from Explore Bristol Research,
<http://research-information.bristol.ac.uk>**

Author:

Jakimowicz, Agata K

Title:

Towards osteogenesis

utilising the power of cell-free protein synthesis for regenerative medicine

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>. This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited in Explore Bristol Research. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.

Towards Osteogenesis: Utilising the Power of Cell-Free Protein Synthesis for Regenerative Medicine

By

AGATA K. JAKIMOWICZ



University of
BRISTOL

School of Cellular and Molecular Medicine

A dissertation submitted to the University of Bristol in accordance with the
requirements of the degree of
DOCTOR OF PHILOSOPHY
in the Faculty of Life Sciences

August 2022

Word count: 41,204

ABSTRACT

The annual incidence of bone disorders and breakages is on the rise due to an ageing population. Numerous drawbacks and constraints of the current clinical treatments for bone repair have been observed, with current demand for bone grafts outstripping supply from donors. Recently, interest in bone regeneration has shifted to the field of tissue engineering, which uses techniques to generate new tissue *in vitro*, such as engineered grafts for implementation, or *in vivo*, by prompting the tissue to self-repair through the use of biomolecules such as growth factors. Here, the latter approach of tissue engineering is combined with synthetic biology to provide the foundations for generating a growth-factor-producing gel capable of stimulating osteogenic differentiation of mesenchymal stem cells (MSCs). The component that enables programmable production of the growth factor, specifically bone morphogenetic protein 2 (BMP2), is created by a cell-free expression system (CFES), an *in vitro* protein synthesis approach. In this project, a collection of growth factor-reporter fusion constructs were designed and selected for the highest yielding by *E.coli* CFES. Secondly, the effect of the fusion interactions with MSCs was visualised, confirming bioactivity of the novel protein. Phosphorylation of downstream BMP2 signalling molecules was detected from fusion samples, indicating significant deGFP-BMP2 receptor-binding activity. Furthermore, an upregulation of osteogenic markers, calcium deposition and enhanced alkaline phosphatase activity were observed, and in some cases to a greater effect than a positive control, rhBMP2, indicating that the fusion protein is capable of MSCs differentiation towards the osteogenic lineage. Lastly, CFES was incorporated into a series of hydrogel networks, where the increasing fluorescence signalled the successful in-gel production of the fusion protein. Looking ahead, this CFPS system could be developed into an injectable gel capable of spatio-temporal osteogenic stimulation, which would be a powerful technique in tissue engineering. This thesis serves as an integral step towards this goal.

ACKNOWLEDGEMENTS

My gratitude extends to everyone who has helped me and supported me during my PhD. Firstly, I would like to thank my primary supervisor, Professor Adam Perriman, for believing in me and providing me with the opportunity to work on this project. The ever-positive thinking and enthusiasm has always been appreciated. In addition, I would like to thank Dr David Lunn, Dr Sam Olof and everyone else who was involved in OxSyBio for funding this project and providing initial guidance. Dr Ross Anderson, for agreeing to be my secondary supervisor and for allowing me to use the equipment and reagents in the C101 lab. Dr Angelique Coutable, for the expertise in cell-free protein synthesis, guidance and patience in my early PhD journey.

I would like to thank the Bristol Centre for Functional Nanomaterials, especially Professor Annela Seddon and Becky Freshwater. My 2017 BCFN PhD cohort, for an incredible first year, with special thanks to Yushi Yang and Mariam Khalfeiy as my THETA partners.

Huge thank you to every Perriman group member, past and present, especially those during my time in the group: Jenna Shapiro, Corrigan Hicks, Peter Johnson, Beth Hickton, George Klemperer, Runa Begum, Thomas Richardson, Ioannis Zampetakis, Sara Carreira, Andrea Diaz Gaxiola, Mark Shannon, Ximena Vasto Anzaldo, Graham Day, Rosy Cuahtecontzi and Valeria Sandoval Torres. Your friendship and support have been incredible, and you all made my PhD journey unforgettable. Special thanks to the girls for the coffee breaks and amusing chats. William Macalester, thank you for literally never leaving my side, it has been a wild ride. William Zhang and Wenjin Xiao, living and working with you both was so much fun, thank you for that. Raquel Cruz Samperio, I cannot thank you enough for everything you have done for me. From having exciting coffee breaks, guiding, and helping me with some experiments, providing me with immense support, to editing this entire thesis. You have always been there for me, for which I will be eternally grateful for.

William Hoffman and Fabiola Cardoso Delgado, thank you for being the best bubble buddies. All the dinners, movie nights and just hanging out have kept me sane.

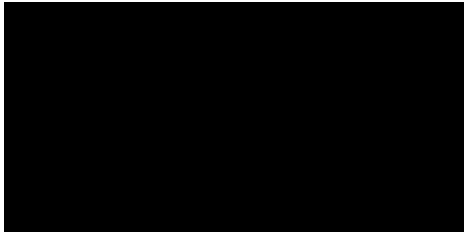
I would like to thank Niall Mulkerns for his immeasurable contribution to my life since 2017. You have always been looking after me, providing me with kind words whenever I needed to hear them the most, and spending your free time on helping with my thesis. I would not have done this without you.

Lastly, huge thank you to my family. Monika, whether it is life or career, you have always been a role model to me, also thank you for reading some of my chapters. Piotr, thank you for putting up with me in the Eastleigh house and looking after me when times were tough. To my parents, for your love and support, for welcoming me into your home for thesis writing and accommodating my every need. You have been absolutely crucial to my success.

AUTHOR'S DECLARATION

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED:



DATE: 08 08 2022

TABLE OF CONTENTS

Abstract	i
Acknowledgements	iii
Author's declaration.....	v
Table of Contents.....	vi
List of Figures	xi
List of Tables.....	xviii
Abbreviations	xix
Chapter 1. Introduction.....	1
1.1 Bone Regeneration	2
1.1.1 Mesenchymal Stem Cells (MSCs)	4
1.1.2 Bone Morphogenetic Proteins (BMPs).....	8
1.1.3 Impaired bone healing and disease.....	12
1.1.4 Clinical standards and tissue engineering.....	13
1.2 Cell-Free Protein Synthesis	17
1.1.5 Protein synthesis in E.coli.....	17
1.1.6 History and components of the cell-free protein synthesis	21
1.1.7 Advantages, limitations and applications.....	24
1.3 Project Design and Aims.....	29
1.4 Thesis Structure	31
1.4 Bibliography	32

Chapter 2. Plasmid Design	48
2.1 Introduction	49
2.2 Materials and Methods.....	52
2.2.1 Gateway cloning	52
2.2.1.1 Creating VEGF and BMP2 sequences.....	53
2.2.1.2 BP reaction	54
2.2.1.3 LR reaction	55
2.2.2 Engineering a two-plasmid system.....	55
2.2.2.1 Restriction digestion.....	55
2.2.2.2 Gel electrophoresis.....	56
2.2.2.3 Ligation and transformation.....	56
2.2.3 Gibson assembly	57
2.2.3.1 Generation of linear DNA fragments.....	58
2.2.3.2 Gibson assembly technique.....	58
2.2.3.3 Colony Polymerase Chain Reaction	59
2.3 Results and Discussion	61
2.3.1 pCellFree	61
2.3.2 Two-plasmid system	63
2.3.3 deGFP-BMP2.....	66
2.4 Conclusions and Further work	69
2.5 Bibliography.....	70

Chapter 3. Optimisation of the Cell-Free Expression System ...	73
3.1 Introduction	74
3.2 Materials and Methods.....	77
3.2.1 In-house cell-free expression system.....	77
3.2.1.1 Amino acid solution	78
3.2.1.2 Energy solution	78
3.2.1.3 Transcription-Translation (TX-TL) reaction setup.....	79
3.2.2 PURExpress reactions setup.....	79
3.2.3 <i>E.coli</i> S30 extract reactions setup.....	80
3.2.4 Hydrogel chassis.....	80
3.3 Results and Discussion	82
3.3.1 deGFP	82
3.3.2 pCellFree	84
3.3.3 Two plasmid system	95
3.3.4 deGFP-BMP2.....	98
3.3.5 Cell-free protein synthesis in hydrogels.....	104
3.4 Conclusions and Further work	115
3.5 Bibliography	119
Chapter 4. Characterisation of the deGFP-BMP2 Chimera	125
4.1 Introduction	126
4.2 Materials and Methods.....	131
4.2.1 Small scale expression of deGFP-BMP2 in <i>E.coli</i>	131

4.2.2	Large scale expression of deGFP-BMP2 in <i>E.coli</i>	132
4.2.3	Purification of deGFP-BMP2.....	133
4.2.3.1	Immobilised Metal Affinity Chromatography (IMAC)	133
4.2.3.2	Size Exclusion Chromatography (SEC).....	134
4.2.4	Polyacrylamide gel electrophoresis (PAGE)	135
4.2.5	Circular Dichroism.....	136
4.2.6	Ultraviolet-visible spectroscopy.....	137
4.2.7	Fluorescence spectroscopy.....	138
4.2.8	Dynamic Light scattering.....	139
4.3	Results and Discussion	140
4.3.1	Expression of deGFP-BMP2	140
4.3.2	Secondary structure analysis of chimera	146
4.3.3	Tertiary structure analysis of chimera	150
4.4	Conclusions and Further work	155
4.5	Bibliography	157
Chapter 5. Differentiation of Human Mesenchymal Stem Cells Using deGFP-BMP2		163
5.1	Introduction	164
5.2	Materials and Methods.....	167
5.2.1	Cell culture	167
5.2.2	Confocal microscopy	168
5.2.3	Flow cytometry.....	169
5.2.4	Western blotting.....	170

5.2.5	Alkaline phosphatase assay	172
5.2.6	Reverse Transcription Quantitative Polymerase Chain Reaction (RT-QPCR) 173	
5.2.7	Alizarin Red S staining	175
5.2.8	Statistical analysis	176
5.3	Results and Discussion	177
5.3.1	Visualising deGFP-BMP2 chimera	177
5.3.2	Downstream signalling of deGFP-BMP2	184
5.3.3	Differentiation assays.....	186
5.4	Conclusions and Further work	193
5.5	Bibliography	195
Chapter 6. Conclusions and Future Project Outlook		199
6.1	Overview	199
6.1	Bibliography	204
A.	Appendix A.....	205
B.	Appendix B.....	217
C.	Appendix C.....	218

LIST OF FIGURES

Figure 1.1 Schematic representation of the bone healing process.	3
Figure 1.2 Schematic of the human mesenchymal stem cells (hMSCs) multilineage.....	4
Figure 1.3 Osteoblast formation flowchart.....	6
Figure 1.4 Osteoblast maturation progression.	7
Figure 1.5 Schematic of the signalling pathways involved in osteoblast homeostasis..	8
Figure 1.6 The subgroups of BMPs and their relationship with the cell surface receptors they bind to, as well as the Smad proteins they activate.	10
Figure 1.7 Bone morphogenetic protein 2 as A: pre-proprotein and B: dimer generated from the mature region (green and blue) bound to extracellular domains of BMPRIA and ActRIIB receptors, viewed at two different angles.	11
Figure 1.8 Schematic of the bone regeneration treatment options.....	15
Figure 1.9 Schematic of the prokaryotic transcription by RNA polymerase.	18
Figure 1.10 Schematic of the process of prokaryotic translation.....	20
Figure 1.11 A basic coupled transcription-translation reaction set-up.....	22
Figure 1.12 A repertoire of applications made possible from cell-free protein synthesis..	26
Figure 2.1 Schematic diagram showing gateway cloning using A: BP reactions and B: LR reactions. R= recombination event between corresponding att sites.	53
Figure 2.2 Agarose gel electrophoresis of PCR products.....	61
Figure 2.3 Agarose gel electrophoresis of pCellFree and Gateway cloned pCellFree plasmids.....	62
Figure 2.4 Agarose gel electrophoresis after restriction digestion of BMP2 with ApaI and KpnI HF restriction enzymes.....	64

Figure 2.5 Agarose gel electrophoresis after vector digestion and ligation with BMP2 insert.	65
Figure 2.6 Agarose gel electrophoresis of PCR reactions for the amplification of Gibson fragments.....	66
Figure 2.7 Agarose gel electrophoresis following colony PCR of Gibson assembled N-terminal deGFP-BMP2.	67
Figure 3.1. Visual quality assessment of viable extract.	78
Figure 3.2 Calibration of extracts 1-4 (A-D) with Mg-glutamate (i) and K-glutamate (ii).	83
Figure 3.3 Kinetics of deGFP (black), eGFP (red and blue), mCherry (green and purple), eGFP-BMP2 fusion (yellow and cyan) and mCherry-VEGF fusion (brown and olive) expression from 9 nM input plasmids.	84
Figure 3.4 Kinetics of eGFP-BMP2 expression in the in-house CFES supplemented by 0 U T7 RNAP (black), 10 U T7 RNAP (red), 30 U T7 RNAP (blue) and 60 U T7 RNAP (green).	86
Figure 3.5 Heatmap of the addition of IPTG and T7 RNAP at a range of concentrations to CFES containing 15 nM pCellFreeG03+BMP2.	88
Figure 3.6 Maximum fluorescence intensity over a 16 hours time-lapse comparing deGFP and pCellFreeG03 expression from in-house or PURExpress CFES.	89
Figure 3.7 Kinetics of eGFP (pCF3 and pCF4), mCherry (pCF5 and pCF6) and their fusions with BMP2 or VEGF expression from 15 nM input plasmids with PURExpress.	91
Figure 3.8 Maximum fluorescence intensity over a 16 hour time-lapse. Reactions of input DNA with three different sources of CFES, S30 extract-based T7 Promega kit (solid colour), in-house CFES (horizontal lines), and PURExpress (diagonal lines).....	93

Figure 3.9 Kinetics of eGFP-BMP2 over a 16-hour timelapse, with Pr1T7 and pCF3+BMP2 at various concentrations as input DNA with in-house CFES.	94
Figure 3.10 Kinetics of deGFP (black) AqpZ-eGFP (colour) over a 16-hour timelapse, with 1 nM pBEST-OR2-OR1-Pr-UTR1-deGFP-T500, and pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 with pTar-AqpZ at various concentrations as input DNA with in-house CFES.....	96
Figure 3.11 Kinetics of deGFP (black) and BMP2-eGFP (colours) over a 16-hour timelapse.	97
Figure 3.12 Maximum fluorescence intensity of deGFP-BMP2 over a 16-hour time-lapse. 1 nM and 5 nM input DNA originating from colonies numbered 1, 5, 8 and 10.....	98
Figure 3.13 Maximum fluorescence intensity of deGFP and deGFP-BMP2 over a 16-hour time-lapse, with various input DNA concentrations.	99
Figure 3.14 Maximum fluorescence intensity of deGFP-BMP2 over a 16-hour time-lapse, with the CFES reaction containing various input DNA concentrations, and carried out at 29°C and 37°C..	101
Figure 3.15 Kinetics of deGFP-BMP2 chimera produced from three extracts over a 16-hour timelapse..	102
Figure 3.16 Standard curve of deGFP-BMP2 in extract 4 CFES.	103
Figure 3.17 Kinetics of deGFP produced from bioink components containing CFES... ..	105
Figure 3.18 Kinetics of deGFP produced from bioink containing CFES, over a 21-hour time-lapse.	106
Figure 3.19 Kinetics of deGFP produced from agar and agarose containing CFES, over a 22-hour time-lapse..	107

Figure 3.20 Diffusion monitoring set-up in a 96-well plate where the gel was pipetted to one side of the well and its fluorescence was measured at Position 1, and once the gel was set it was submerged in DEPC water where fluorescence was measured at Position 2. .	108
Figure 3.21 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area where the agar or agarose gels were placed in a 96-well plate (Position 1).	110
Figure 3.22 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area furthest away from where the agar or agarose gels were placed in a 96-well plate (Position 2)..	111
Figure 3.23 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area where the bioink gels were placed in a 96-well plate (Position 1).	113
Figure 3.24 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area furthest away from where the bioink gels were placed in a 96-well plate (Position 2).	114
Figure 4.1 Schematic representation of the hypothesised deGFP at three various angles.	127
Figure 4.2 Ribbon representation of BMP2 dimer formed from two monomers labelled in green and blue.	128
Figure 4.3 Schematic representation of the hypothesised deGFP-BMP2 chimera at three various angles.	129
Figure 4.4 Reduced SDS PAGE of small-scale deGFP-BMP2 expression in Rosetta, SHuffle and BL21 bacterial cells.	140
Figure 4.5 Reduced SDS PAGE of small-scale expression of deGFP-BMP2 in Rosetta cells taken place over 24 hours or 48 hours at 25°C or 37°C.	141
Figure 4.6 IMAC purification of deGFP-BMP2.....	143
Figure 4.7 SEC purification of deGFP-BMP2.	144

Figure 4.8 Lane content analysis from Figure 3.4 B SDS PAGE, performed by GelAnalyzer.....	146
Figure 4.9 Circular dichroism (CD) spectra of A: EGFP and B: deGFP-BMP2 chimera.	147
Figure 4.10 Fitted circular dichroism (CD) spectra obtained at a temperature gradient with A: EGFP and B: deGFP-BMP2 chimera.....	149
Figure 4.11 UV-vis spectrum of EGFP (red) and deGFP-BMP2 chimera (black).....	150
Figure 4.12 Fluorescence excitation (solid lines) and emission (dashed lines) spectrum of EGFP (red) and deGFP-BMP2 chimera (black).	152
Figure 4.13 Dynamic light scattering (DLS) measurements to determine the hydrodynamic diameter of EGFP (red) and deGFP-BMP2 (black).....	153
Figure 5.1 Schematic representation of the BMP2 signalling cascades including the canonical Smad and non-canonical pathways.	165
Figure 5.2 Z-stack confocal images of hMSCs incubated with 400 µg/mL <i>E.coli</i> -produced deGFP-BMP2 chimera for three different time periods of A: 20 minutes, B: 80 minutes and C:120 minutes.....	178
Figure 5.3 Flow cytometry gating strategy displaying density plots of 10,000 events (low density blue to high density red). F	181
Figure 5.4 Flow cytometry histograms of GFP fluorescence presented as FITC-A GFP vs cell count for each tested condition.	182
Figure 5.5 Western blot assessing the phosphorylation of Smad proteins 1 and 5 in cell lysates when hMSCs were incubated for 4 hours with 25 ng/mL rhBMP2 (positive control), 25 ng/mL and 200 ng/mL <i>E.coli</i> -produced deGFP-BMP2 chimera, and complete 10 µL CFES reactions of deGFP-BMP2 chimera and deGFP per 1 mL of media.	185

Figure 5.6 ALP activity measured from lysates originating from hMSCs incubated for 7 days in either expansion media denoted as ‘E’ or osteogenic media denoted as ‘O’, supplemented with 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2, or 0 ng/mL BMP2.187

Figure 5.7 Quantitative PCR of *Runx2*, *Col1A1*, *ALP* and *BGLAP* from hMSCs conditioned for 14 days in osteogenic media containing 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2.188

Figure 5.8 Representative Alizarin Red S stain images (A-E) and absorbance scan (F) from hMSCs cultured for 28 days in osteogenic differentiation media.191

Figure 5.9 Alizarin Red S quantified from each tested condition after the stain was extracted and standard curve was performed.192

Figure A.1 Maps of pCellFreeG03 plasmids. A: Unmodified pCellFreeG03 consisting of EGFP with 8xHis tag at the N-terminal and HRV-3C site along with attB sites at the C-terminal. B: Gateway cloning modified pCellFreeG03 consisting of 8xHis tag, EGFP, HRV-3C site, full length BMP2, respectively.....207

Figure A.2 Maps of pCellFreeG04 plasmids. A: Unmodified pCellFreeG04 consisting of EGFP with 8xHis tag at the C-terminal and HRV-3C site along with attB sites at the N-terminal. B: Gateway cloning modified pCellFreeG04 consisting of full length BMP2, HRV-3C site, EGFP and 8xHis tag, respectively.....208

Figure A.3 Maps of pCellFreeG05 plasmids. A: Unmodified pCellFreeG05 consisting of mCherry with 8xHis tag at the N-terminal and HRV-3C site along with attB sites at the C-terminal. B: Gateway cloning modified pCellFreeG05 consisting of 8xHis tag, mCherry, HRV-3C site, VEGF, respectively.....209

Figure A.4 Maps of pCellFreeG06 plasmids. A: Unmodified pCellFreeG06 consisting of mCherry with 8xHis tag at the C-terminal and HRV-3C site along with attB sites at the N-terminal. B: Gateway cloning modified pCellFreeG06 consisting of VEGF, HRV-3C site, mCherry and 8xHis tag, respectively.....	210
Figure A.5 Map of Pr1-T7 plasmid. Encodes for a T7 RNA polymerase. It was used in conjunction with pCellFree plasmids to bring about their transcription.....	211
Figure A.6 Map of sigma 28 plasmid. Once expressed, it enables transcription of plasmids with pTar promoters.....	212
Figure A.7 Map of unmodified expression plasmid. Requires sigma factor 28 to be transcribed. Encodes for AqpZ with EGFP at the C-terminal. This plasmid was used to replace AqpZ gene with mature-length BMP2 gene.....	213
Figure A.8 Map of modified expression plasmid encoding for mature-length BMP2, TEV cleavage site and EGFP, respectively. Requires sigma factor 28 to be transcribed.....	214
Figure A.9 Map of pBEST-OR2-OR1-Pr-UTR1-deGFP-T500.....	215
Figure A.10 Map of modified pBEST-OR2-OR1-Pr-UTR1-deGFP-T500, now encoding for deGFP, TEV cleavage site, mature-length BMP2, 6xHis tag, respectively.....	216
Figure B.1 Naturally occurring amino acids. Generated from biorender.com.....	217

LIST OF TABLES

Table 1.1 A list of some components that are added to certain cell extracts, and their roles in increasing the yield of the synthesised proteins.....	23
Table 2.1 Thermocycler programme for generating linear VEGF and BMP2 sequences for Gateway cloning.....	54
Table 2.2 Thermocycler programme for generating linear DNA fragments.....	58
Table 2.3 Thermocycler programme for Colony PCR.....	59
Table 2.4 Summary of the plasmids obtained and generated in this thesis.....	60
Table 3.1 Average concentrations of deGFP-BMP2 from in-house CFES reactions carried out at 29°C with different input DNA concentrations.....	103
Table 4.1 CD measurement settings.....	137
Table 4.2 UV-vis measurement settings.....	137
Table 4.3 Fluorometry measurement settings.....	138
Table 5.1 Components of cell culture media used in the assays detailed below.....	168
Table 5.2 Summary of flow cytometry results of GFP positive populations, expressed as percentages for each repeat per condition.....	183
Table 5.3 Quantification of the Western blot using GelAnalyzer software.....	185

ABBREVIATIONS

ALP	alkaline phosphatase
AqpZ	aquaporin z
ARS	alizarin red s
att	site-specific-attachment
BMPRI	bone morphogenetic protein type I receptor
BMPRII	bone morphogenetic protein type II receptor
BMPs	bone morphogenetic proteins
BSP	bone sialoprotein
CD	circular dichroism
cDNA	complementary deoxyribonucleic acid
CFES	cell-free expression system
CFPS	cell-free protein synthesis
Col1A1	collagen type I alpha I
Co-Smads	common-partner Smads
deGFP	enhanced green fluorescent protein as a minimal domain with amino acids 6-229
DEPC	diethylpyrocarbonate
DLS	dynamic light scattering
DLX5	drosophila distal-less 5
DMEM	dulbecco's modified eagle medium
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E.coli	escherichia coli
ECM	extracellular matrix
EDTA	ethylendiaminetetraacetic acid
EF	elongation factor

EGFP	enhanced green fluorescent protein
FACS	fluorescence-activated cell sorting
FBS	foetal bovine serum
FDA	US Food and Drug Administration
FGF	fibroblast growth factor
FSC-A	forward scatter area
FSC-H	forward scatter height
hMSCs	human mesenchymal stem cells
HRP	horseradish peroxidase
IHF	integration host factor proteins
IL	interleukin
IMAC	immobilised metal affinity chromatography
Int	integrase
IPTG	isopropyl- β -d-thiogalactopyranoside
Lac	lactose
LB	luria broth
LDS	lithium dodecyl sulphate
mRNA	messenger ribonucleic acid
MSCs	mesenchymal stem cells
MUP	4-methylumbelliferyl phosphate disodium salt
MWCO	molecular weight cut-off
NTA	nitrilotriacetic acid
NTPs	ribonucleoside 5'-triphosphates
OCN/BGLAP	osteocalcin
OPN	osteopontin
ORF	open reading frame
OSX	osterix

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
pCF	pCellFree
PCR	polymerase chain reaction
pdeGFP	pBEST-OR2-OR1-Pr-UTR1-deGFP-T500
pdeGFP-BMP2	pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500
PDGF	platelet-derived growth factor
PDMS	polydimethylsiloxane
PGA	polyglycolic acid
PLA	polylactic acid
PLGA	poly(lactic-co glycolic acid)
PMSF	phylmethylsulfonyl fluoride
PTH	parathyroid hormone
PVDF	polyvinylidene fluoride
RBS	ribosome-binding site
RFU	relative fluorescence units
rhBMP2	recombinant human bone morphogenetic protein 2
RNAP	ribonucleic acid polymerase
RNases	ribonucleases
R-Smads	receptor-regulated Smads
RUNX2	runt-related transcription factor 2
SDS	sodium dodecyl sulphate
SEC	size exclusion chromatography
SITS	species independent translation initiation sequence
SSC-A	side scatter area
TBS	tris-buffered saline
TGF- β	transforming growth factor beta

TNF- α	tumour necrosis factor alpha
TrB	Transfer buffer
tRNAs	transfer ribonucleic acids
TX-TL	transcription-translation
UV-Vis	ultraviolet-visible
VEGF	vascular endothelial growth factor
YENB	yeast extract nutrient broth
(q)PCR	(quantitative) polymerase chain reaction
2xYT+P	yeast extract with tryptone and phosphates
3-PGA	3-phosphoglyceric acid

CHAPTER 1. INTRODUCTION

1.1 BONE REGENERATION

Bone remodelling occurs throughout one's lifetime in response to mechanical and metabolic requirements, and the process is highly regulated by cytokines and growth factors such as bone morphogenetic proteins (BMPs). These biomolecules orchestrate the interplay between anabolic osteoblasts, which form the bone and catabolic osteoclasts, which contribute to bone resorption. The balance can sometimes become impaired in delayed union or non-union fractures, as well as disease states such as osteoporosis. It is estimated that for approximately 10% of all fractures, the healing process is impaired, which can lead to patient morbidity and chronic pain¹. In those cases where the requirements for an effective bone regeneration surpass the ability of self-healing, clinical interventions are necessary.

Studies suggest that bone repair process closely resembles that of skeletal formation during embryonic development, where both are tightly temporally and spatially controlled by the same biomolecules². The process of bone fracture healing consists of several overlapping stages and can be divided into an inflammatory phase, renewal phase, and a remodelling phase (Figure 1.1). An injury immediately results in an inflammatory response with a release of several inflammatory cytokines such as interleukin-1 (IL-1), IL-6, IL-11 and IL-18, as well as tumour necrosis factor- α (TNF- α)³. This release cascade of inflammatory mediators leads to formation of a clot, known as fracture haematoma due to increased angiogenesis⁴. Nearby platelets release transforming growth factor (TGF- β) and, platelet-derived growth factor (PDGF). The fibrin network created by the haematoma allows for the influx of inflammatory cells which are attracted to the signals released by damaged cells and extracellular matrix, as well as local macrophages, also known as osteomacs⁵. Macrophages attracted by the neutrophils remove the fibrin network along with necrotic cells and secrete further inflammatory mediators and chemokines. These signals initiate the arrival of mesenchymal stem cells (MSCs), fibroblasts and osteoprogenitor cells to the

site of the fracture⁶. Damaged bone fragments are removed by osteoclasts *via* granular breakdown.

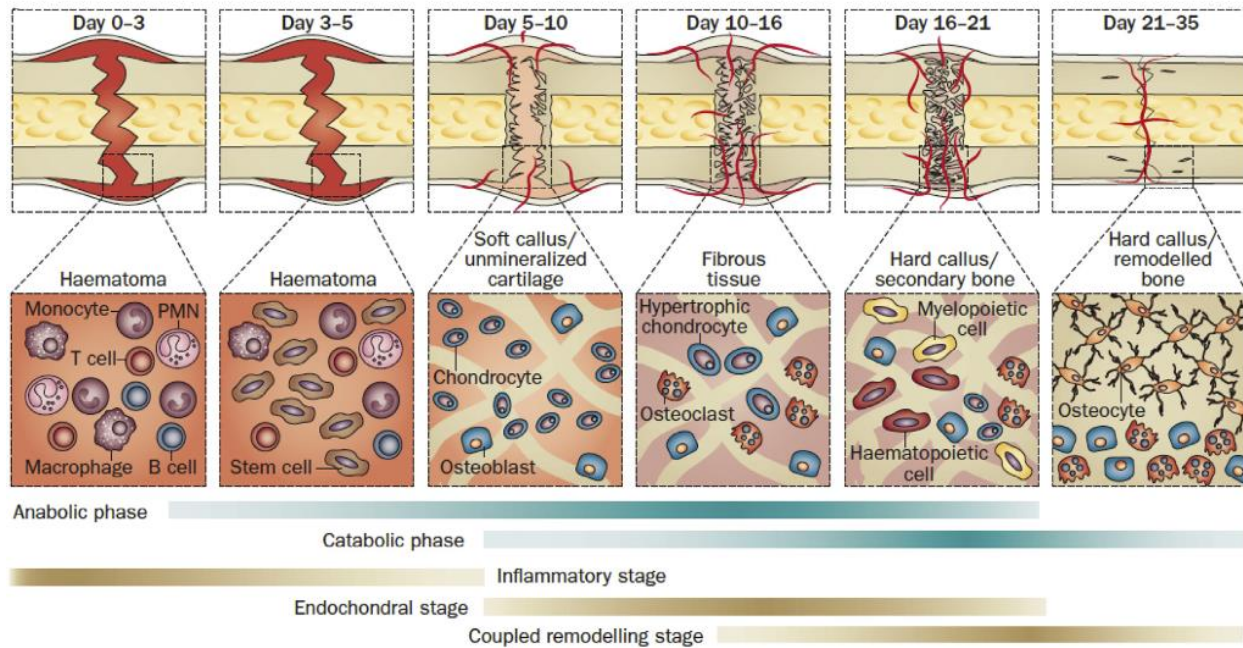


Figure 1.1 Schematic representation of the bone healing process. A haematoma is formed in the first few days of the bone breakage, followed by a soft callus that turns into a fibrous tissue halfway through the healing process. Hard callus is a term for the last phase, where the calcified tissue is remodelled by osteoclasts. Figure adapted from Einhorn and Gerstenfeld⁷.

The osteoprogenitor cells and MSCs proliferate and differentiate, and these processes are guided by the growth factors and mediators present at the site. At this stage, TGF- β , PDGF, vascular endothelial growth factor (VEGF), fibroblast growth factor (FGF) and bone morphogenetic proteins (BMPs) are the key growth factors associated with the formation of granulation tissue which is full of differentiating MSCs *via* the chondrogenic pathway and growing new vasculature⁸. This soft callus acts as a first mechanical support onto which endochondral bone formation will take place. The cartilage produced by chondrocytes will eventually connect the two broken ends of the fracture. As this soft callus develops, the recruited MSCs differentiate into osteoblasts that directly form the immature woven bone with a disorganised lamellar structure that is deposited on the cartilage scaffold⁹. The chondrocytes will undergo apoptosis, leaving behind a calcified

extracellular matrix *via* calcium secretion. The hard callus stage is also a time where vascular ingrowth is enhanced, increasing the blood flow into the site of repair¹⁰. The final stage of bone repair is the remodelling, which is carried out by osteoclasts to restore the original form of the bone¹¹. It is initiated by the osteoblasts expressing IL-1, IL-6 and IL-11 to form osteoclasts. Other factors including TNF- α , IL-12, and interferon- γ are also regulated during the remodelling phase. Due to the opposing roles of osteoblasts and osteoclasts, the woven bone is replaced by the lamellar bone¹². Bone turnover is heightened usually six months post fracture and does not lower to baseline for numerous years after¹³.

1.1.1 MESENCHYMAL STEM CELLS (MSCs)

MSCs play a leading role in the generation of bone forming cells, and therefore are crucial in bone regeneration. Human bone marrow is one of the primary sources of multipotent progenitor cells, also known as MSCs, that self-renew and can undergo differentiation into distinctive cell lineages, including osteoblasts, chondrocytes, fibroblasts, adipocytes, and other tissue-forming cells (Figure 1.2)¹⁴. The multipotency can be characterised by the stem cell's ability to proliferate whilst plastic-adherent *in vitro*, have consistent surface markers whilst multipotent, and be capable of multilineage differentiation by specific molecules in the environment *in vitro*¹⁵.

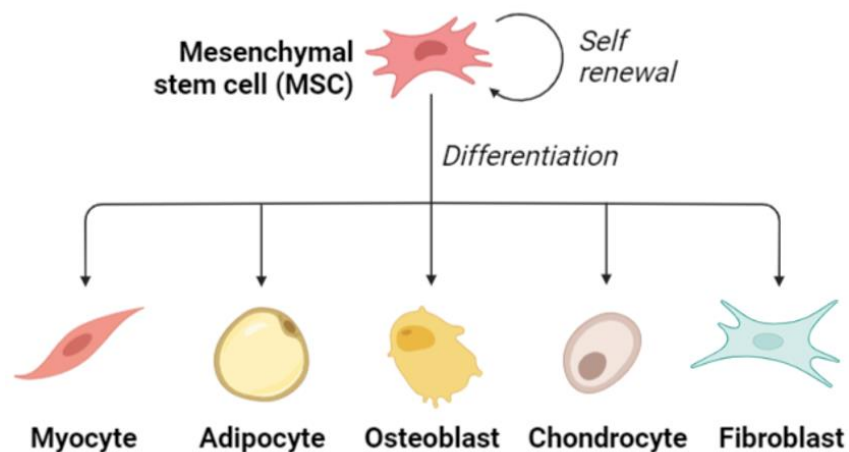


Figure 1.2 Schematic of the human mesenchymal stem cells (hMSCs) multilineage. MSCs can self-replicate or undergo multilineage differentiation. Figure generated using biorender.com.

Whilst MSCs remain multipotent, the surface markers by which they are identified by include CD29, CD44, CD71, CD73 CD90, CD105, and CD106, whereas the identifiable surface markers that are downregulated include CD14, CD34, and CD45²⁸. The change in surface markers occurs once MSCs undergo lineage differentiation. Although bone marrow is the major source of the rare MSCs, there are other locations that have now been identified as MSCs niches, including dental pulp¹⁶, endometrium¹⁷, placenta¹⁸ and umbilical cord¹⁹ amongst others. Adult stem cells including haematopoietic cells and mesenchymal cells are quiescent until they are recruited to partake in tissue regeneration arising from disease, injury, or turnover²⁰. The properties of MSCs, including their multipotency as well as paracrine signalling, make them a useful candidate in regenerative medicine. MSCs are currently employed in stem cell therapies in regenerative medicine, involving an injection of the stem cells directly into a site that is damaged. Stem cell therapy has been explored for numerous applications including cartilage²¹, myocardial²² and orthopaedics repair²³. The ability to undergo multilineage differentiation is not the only mechanism by which MSCs contribute to the therapeutic effect; they also contribute through their paracrine factors²⁴ and by secretion of extracellular vesicles²⁵. However, following MSC injections, their poor retention, attachment, and relatively short life contribute to a large barrier for clinical translation²⁶. Alternatively, endogenous MSCs can be recruited from their niche to the damage site for regeneration²⁷. This strategy implements the delivery of biomolecules such as growth factors, chemokines, or cytokines to drive stem cell migration from the niche and differentiation.

There are two pathways by which osteoblasts can arise, although, both have mesenchymal origin; either from mesenchymal progenitors or from osteochondro-progenitors (Figure 1.3). Both methods generate a common preosteoblast (also referred to as osteoblast progenitor). The process of osteoblast maturation from osteoblast progenitor takes place in three stages: proliferation, maturation, and mineralisation.

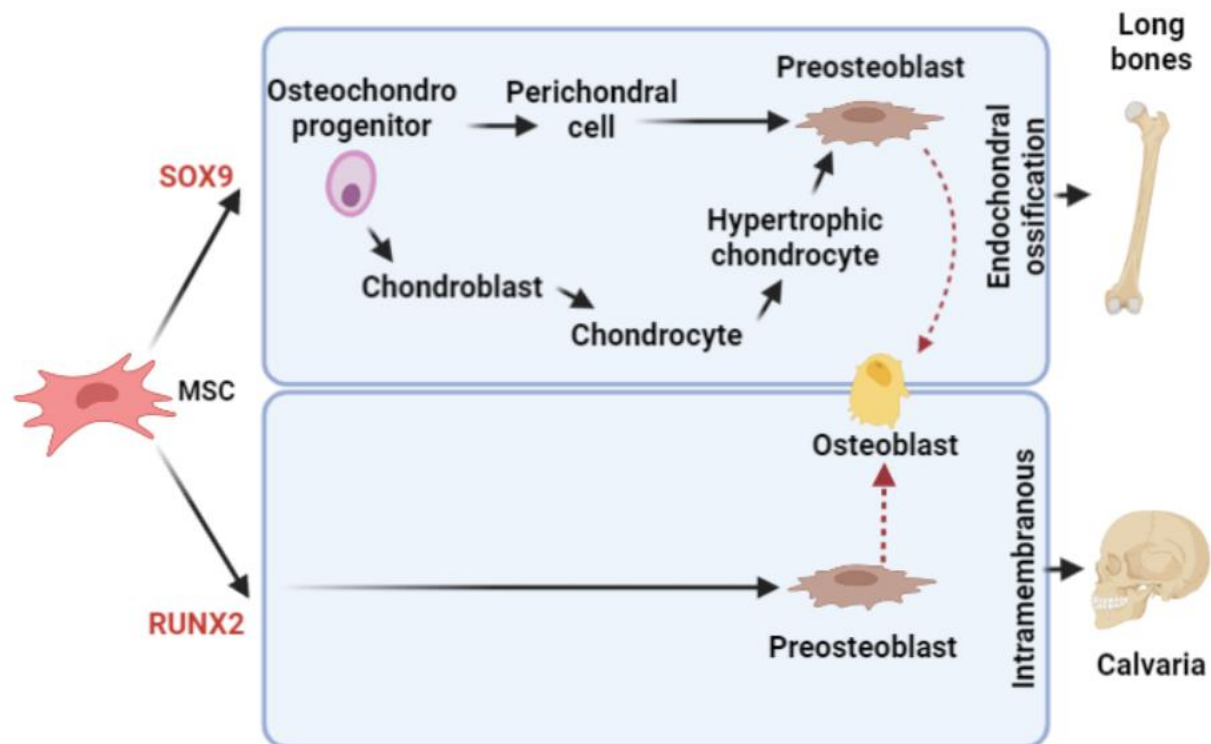


Figure 1.3 Osteoblast formation flowchart. Mesenchymal stem cells (MSCs) express SOX9 for chondrocytic lineage, and RUNX2 for osteoblastic lineage. In intramembranous ossification, preosteoblasts are directly generated from MSCs, whereas in endochondral ossification, they are eventually created from osteochondro progenitors. The osteoblast maturation (red dotted arrows) occurs in three stages: proliferation, maturation, and mineralisation. Figure created in biorender.com.

In the first stage, the cells continue to actively proliferate, which stops by the second stage when the immature osteoblasts express collagen type 1 alpha 1 chain (*Col1A1*) for extracellular matrix (ECM) formation, alkaline phosphatase (*ALP*) for ECM maturation, as well as bone sialoprotein (*BSP*) (Figure 1.4). When the osteoblast matures an upregulation of several osteogenic markers can be noted. These promote mineral deposition in the maturing matrix, including osteocalcin (*OCN*), osteopontin (*OPN*) and *BSP*, with *ALP* and *Col1A1* continuing ECM maturation²⁹.

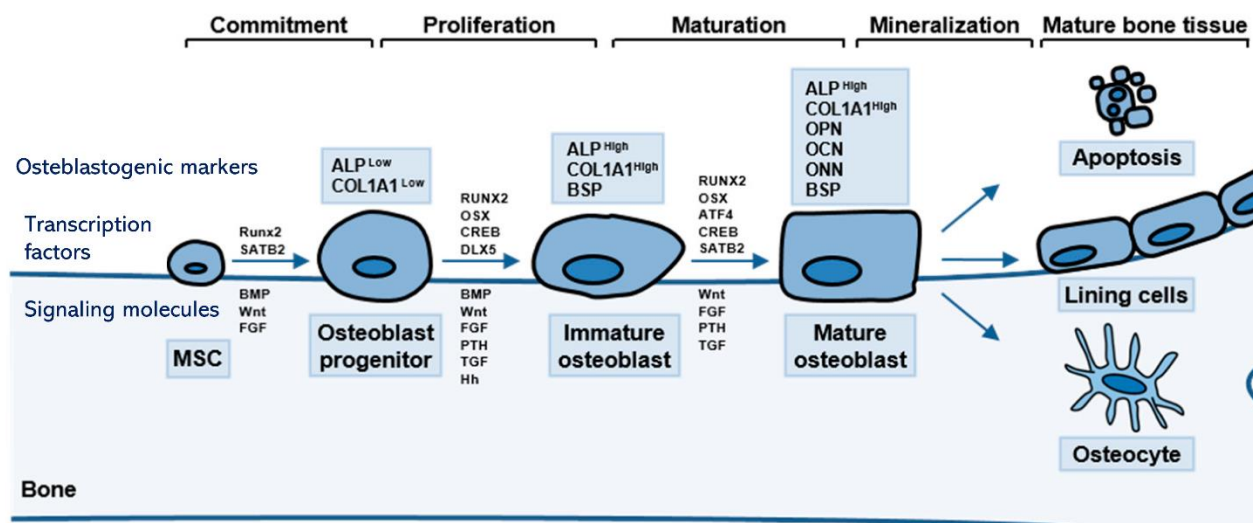


Figure 1.4 Osteoblast maturation progression. Mesenchymal stem cells (MSCs) commit to osteogenic differentiation upon signalling from BMP2, Wnt or FGF pathways, generating the production of a master regulator, RUNX2. With the progression *via* proliferation, maturation and mineralisation in the osteogenic lineage, various osteogenic markers become regulated. Figure adapted from Amarasekara et al.²⁹.

These changes during osteoblast differentiation are initially brought about by extracellular molecules that bind to the surface receptors on cells. The binding event induces intracellular signalling, resulting in the activation and expression of target genes responsible for proliferation, migration, or differentiation³⁰. There are multiple signalling cascades that bring about the generation of new osteoblasts in new bone growth or remodelling, including wntless-related integration site (Wnt), parathyroid hormone (PTH), BMP, TGF- β , FGF, and hedgehog (Hh) (Figure 1.5). Although each of these pathways play an important role in osteoblastogenesis, Wnt and BMP are the most crucial for committing MSCs into osteogenesis differentiation, with BMP2 being identified as a highly potent osteogenic factor^{31, 32}.

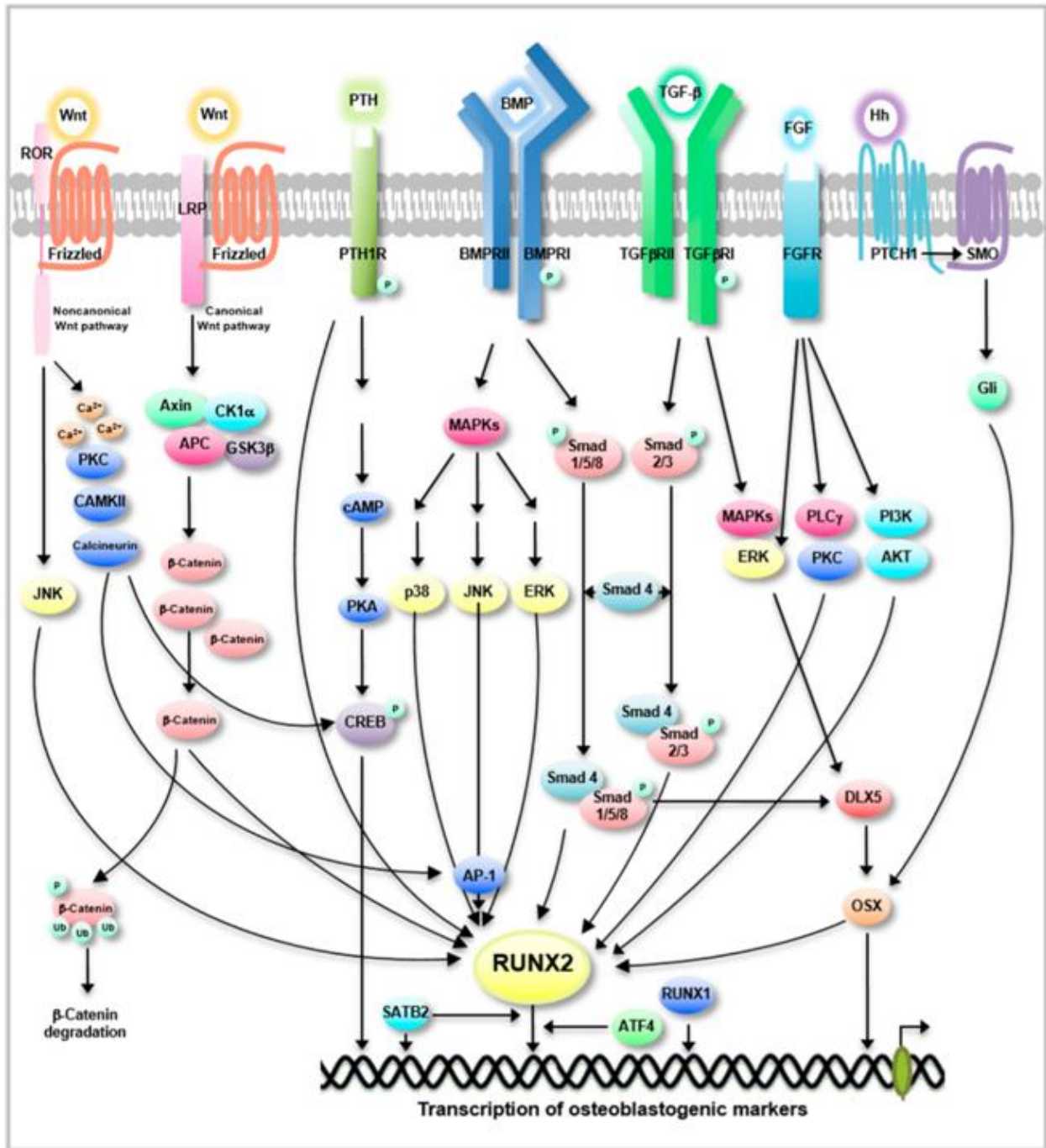


Figure 1.5 Schematic of the signalling pathways involved in osteoblast homeostasis. Wnt, PTH, BMP, TGF- β , FGF and Hh, canonical and non-canonical networks regulating the transcription of osteoblastogenic markers. Figure adapted from Amarasekara et al. ²⁹.

1.1.2 BONE MORPHOGENETIC PROTEINS (BMPs)

BMPs are growth factors originating from a large transforming growth factor- β (TGF- β) superfamily. The highly conserved structures of BMPs are shared with other members of

the TGF- β superfamily, and just like other superfamily members, they are synthesised as a pre-propeptide³³. More than 20 different BMPs have been identified which carry out a biological activity in not only bone, but other tissues such as skeletal muscle³⁴, adipose³⁵, hair follicles³⁶, kidney³⁷, and blood vessels³⁸. The biological activities that BMPs are regulating in the tissues comprise of lineage commitment, proliferation, differentiation, morphogenesis as well as apoptosis³⁹. The name ‘bone morphogenetic protein’ was originally given to the active factors that were first observed to induce new bone formation in demineralised bone matrix before other functions were identified⁴⁰. It was Wang et al., who successfully isolated and identified the BMP2 activity in bone extracts by trypsin digestion and sequencing⁴¹. The family of BMPs can be divided into four subgroups based on their structural homology: BMP2 with BMP4, BMP5 with BMP6/7/8, BMP9 with BMP10, and the last group as BMP12 with BMP13 and BMP14⁴². BMPs operate *via* autocrine or paracrine regulation and initiate signalling by binding to type I and type II cell surface receptors to trigger downstream sequence of events leading to upregulation of osteogenic genes (see Chapter 5.1)⁴³. BMPs are capable of binding to type I receptors only, but the binding affinity is greatly increased in the presence of a type II receptor⁴⁴. There are a number of combinations of the receptors that the ligands bind to, depending on the type of BMP present (Figure 1.6). This also applies to the type of intracellular Smad protein being phosphorylated by the activated type I receptor. Once receptor-regulated Smads (R-Smads) are phosphorylated at their serine motif, they form a complex with a common partner (Co-Smad) and translocate to the nucleus to regulate transcription of target genes³⁹. The canonical and Smad-independent downstream signalling triggered by BMP2 binding to the cell receptors lead to the expression of master regulators of osteogenic differentiation, namely, runt-related transcription factor 2 (*RUNX2/CBFA1*) and osterix (*OSX*)⁴⁵. These further commit stem cells into osteogenic differentiation.

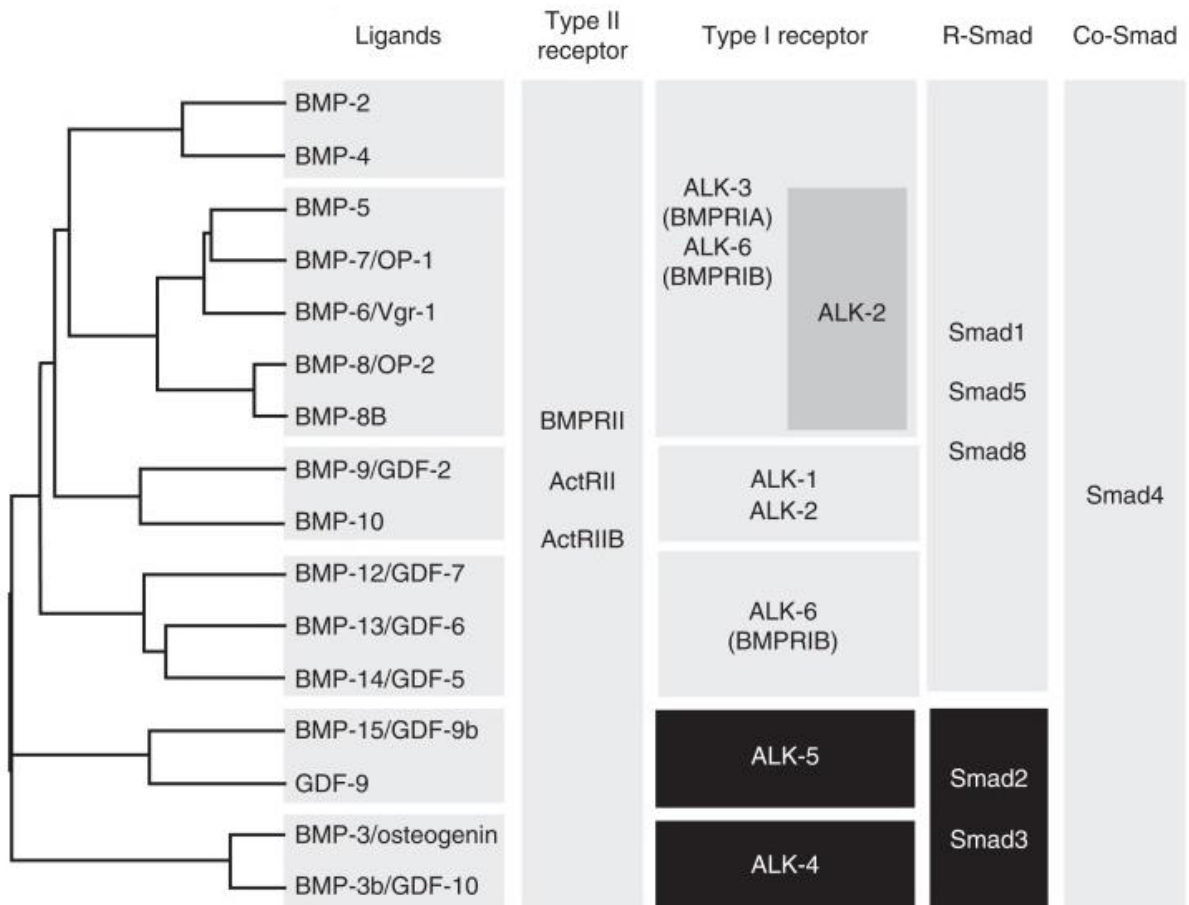


Figure 1.6 The subgroups of BMPs and their relationship with the cell surface receptors they bind to, as well as the Smad proteins they activate. Reproduced from Katagiri and Watabe³⁹.

BMP2 specifically plays a crucial role in bone regeneration, having one of the largest osteogenic capabilities⁴⁶. It is synthesised as a pre-protein which then becomes glycosylated and cleaved to form an active cysteine-knot homodimer from the mature monomers made up of 114 amino acids each (Figure 1.3) (see Chapter 4.1). Upon homodimer formation, the subunits join at a single conserved cysteine residue creating a disulphide bridge⁴⁷. The overall structure of the 26 kDa (32 kDa glycosylated) BMP2 dimer is often referred to as a ‘butterfly’ or ‘fists’, representing its topology. The binding to BMPRI occurs on the wrist epitope of each BMP2 monomer. Once type I receptor binding is complete, type II receptor binding follows by interacting with the knuckle epitopes of the monomers⁴⁸.

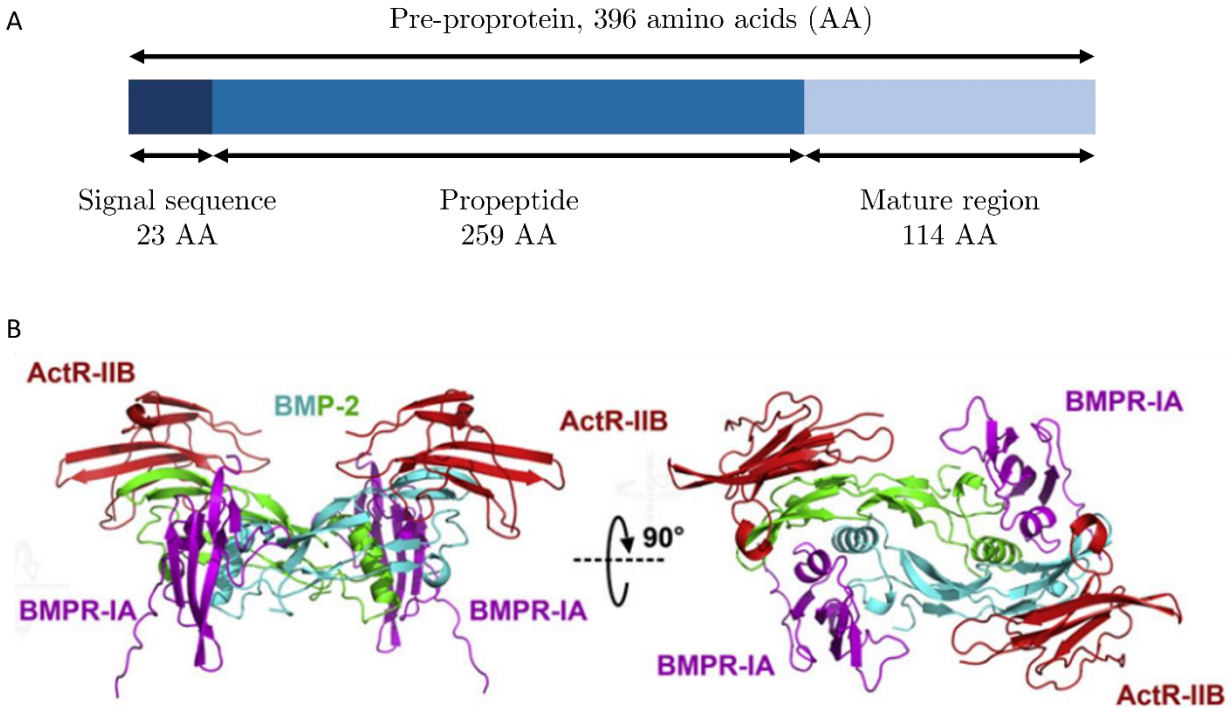


Figure 1.7 Bone morphogenetic protein 2 as A: pre-proprotein and B: dimer generated from the mature region (green and blue) bound to extracellular domains of BMPRII and ActRIIB receptors, viewed at two different angles. Figure B adapted from Mueller and Nickel ⁴⁸.

The bone induction process activated by BMP2 is similar to the endochondral bone formation in embryogenesis, pointing to BMP2 as the main osteogenesis regulator ⁴⁹. Studies on mice with a BMP2 deletion have been conducted and have shown that the resulting embryos were non-viable, whereas conditional knockouts showed a severely altered mechanical properties of bones including reduced strength, radial bone thickness and higher risk of fractures in the underdeveloped bones ^{50, 51, 52}. Additionally, BMP2 knockout mice have presented severe chondrodysplasia⁵³ and altered vasculature architecture ⁵¹. Severe osteoarthritis evidenced by articular cartilage loss and reduced motion range has been reported in mice when BMP signalling was diminished at articular cartilage ⁵⁴.

1.1.3 IMPAIRED BONE HEALING AND DISEASE

Osteoporosis is the most common and prominent bone disease that currently affects more than 200 million people⁵⁵. The numbers are predicted to greatly increase with our aging population. Although osteoporosis can affect any race or sex, it is most prevalent in post-menopausal women and elderly people. In a meta-analysis performed in 2020 and published in the Journal of Orthopaedic Surgery and Research, the prevalence in women was reported at 23.1%, whereas the prevalence of osteoporosis in men was at 11.7% globally⁵⁶. However, the percentages are higher in certain continents and older age groups, and generate a substantial physical, social, and economic burden⁵⁷. The healthcare costs associated with osteoporosis in the European Union in the year 2000 were valued at €32 billion (£21.8 billion) and are predicted to double by year 2050⁵⁸. Osteoporosis is characterised by low bone mass, deterioration of the tissue, and defective bone micro-architecture and its associated extracellular matrix⁵⁹. These characteristics lead to significantly higher risk of bone fractures⁶⁰. Osteoporosis is often referred to as a ‘silent disease’ because it is diagnosed after a fracture occurs. Bone fractures prompt to a chronic pain, morbidity and even mortality, where within one year after a hip fracture, the mortality rate increases by 15-20%⁶⁰. A major factor in arising osteoporosis is the imbalance between bone formation and resorption during remodelling, meaning the activity of osteoblasts is greatly reduced in comparison to the activity of osteoclasts. This imbalance is associated with disrupted BMP2 signalling, where the availability of BMP2 has been shown to be significantly reduced in osteoporotic tissue⁶¹. Mice models have demonstrated that a loss of the *BMP2* gene results in an osteoporosis-like phenotype⁶². Additionally, the activity of *Noggin* and *Gremlin*, both known soluble antagonists of BMP2, has shown to weaken the bones as well as display osteoporosis-like phenotype⁶³.

Aside from osteoporosis, there are numerous diseases that increase the risk of a bone fracture, including osteogenesis imperfecta. Although osteogenesis imperfecta is a genetic disorder where a mutation is present in the chains of collagen type I, the characteristics

encompass bone fragility and low bone mass amongst other abnormalities⁶⁴. Although, depending on the severity of the collagen type I mutations, these manifestations often result in fractures. The second most common bone disease after osteoporosis is Paget disease, where sufferers experience abnormal bone remodelling leading to symptoms such as bone pain, deformities, fracture, as well as hearing loss amongst others⁶⁵. Bone loss can also be attributed to infections, tumours, arthritis, osteonecrosis, and metabolic bone disease, which require medical intervention⁶.

1.1.4 CLINICAL STANDARDS AND TISSUE ENGINEERING

The clinical gold standard for repairing large bone defects is through the use of autografts. This is a process of transferring a healthy piece of bone from donor site, usually the iliac crest, to the defect site within an individual⁶⁶. Autografts are a preferred method of bone repair due to its osteoconductive, osteogenic and osteoinductive properties⁶⁷. This refers to the compatibility of the grafts surface with bone cells, stimulating new bone cells growth, as well as differentiation of cells into osteoblasts, respectively. Due to the graft being of the individual's origin, the risk of autoimmune rejection is minimal. However, the major downfall for this surgery option is the common donor site morbidity, where patients lose the sensitivity and functionality at the harvest site. Other complications near or at the donor site that have been noted include the development of haematomas, hernias, or chronic pain⁶⁸. An alternative method that utilises a bone graft transplant from one individual into another of the same species, is termed allograft. Whereas, xenograft is of another species origin, usually porcine. Both allografts and xenografts allow for larger quantities of the graft, however, the integration time of the graft is much slower and has a lower success rate⁶⁹.

Synthetic bone graft substitutes are continually investigated to advance therapeutic options (Figure 1.8). Calcium phosphate-based materials, such as hydroxyapatite or ceramics are the most common in the production of synthetic grafts⁷⁰. Other synthetic grafts can be made of bioactive glass or calcium sulfate^{71,72}. These synthetic grafts

resemble the mineral phase of the bone, providing an excellent biocompatibility with native bone⁷³. However, their brittleness and low elasticity is a major drawback, which brings the use of polymers into the spotlight⁷⁴. Some of the polymer-based biomaterials can be either natural, such as collagen, alginate and gelatin, or synthetic, including polylactic acid (PLA), polyglycolic acid (PGA) and poly(lactic-co glycolic acid) (PLGA)⁷⁵. Natural polymers are biodegradable and due to their inherent bioactivity, allow for greater interactions with cells when compared to synthetic polymers⁷⁶. Synthetic polymers have proven to be useful in tissue engineering because of their well-established mechanical and physical properties. The methods by which polymers are created and classified are vast, as they are often dictated by the application needs⁷⁷.

In the last two decades, three-dimensional (3D) bioprinting of scaffolds and tissue constructs using cells embedded in biomaterial matrices has greatly advanced the field of tissue engineering and regenerative medicine. Bioprinting allows for the precise deposition control of the functional biochemical cues and cells in the tissue construct, as well as granting the desired final tissue geometry. Cells are usually added to high water content bioinks, most commonly consisting of natural polymers, such as alginate, hyaluronic acid, gelatin or collagen. However, synthetic polymers including polyethylene glycol (PEG), poly(vinyl alcohol) (PVA), or PLGA are also often used⁷⁸. A bioink designed by Armstrong *et al.* has been specifically designed for tissue engineering needs that is compatible for 3D bioprinting⁷⁹. It consists of Pluronic F127 mixed with sodium alginate, and when a desired construct is printed at physiologically-relevant temperature, the Pluronic F127 undergoes a sol-gel transition, solidifying the structure. This is attributed to Pluronic micelle formation at or above the critical micellar concentration, which can be reduced upon temperature elevation. The printed construct is then immersed in calcium chloride which crosslinks the alginate chains, stabilising the templated structure. The immersion of the construct in a liquid that has a temperature below the gelation range, induces the hydration of the hydrophobic poly(propylene oxide) (PPO) units of Pluronic

F127, and both PPO and poly(ethylene oxide) (PEO) segments become soluble, leading to gel-sol transition. The now soluble F127 is removed with subsequent wash steps, leaving behind a highly porous, cross-linked matrix. Armstrong *et al.* was able to show long-term stem cells viability as well as retained ability of MSCs to differentiate into chondrocytes and osteoblasts, within the bioink⁷⁹. As evident, polymers or composites for tissue engineering can be enhanced by cell seeding, however, the addition of growth factors such as BMP2 to obtain osteogenic properties is also achievable.

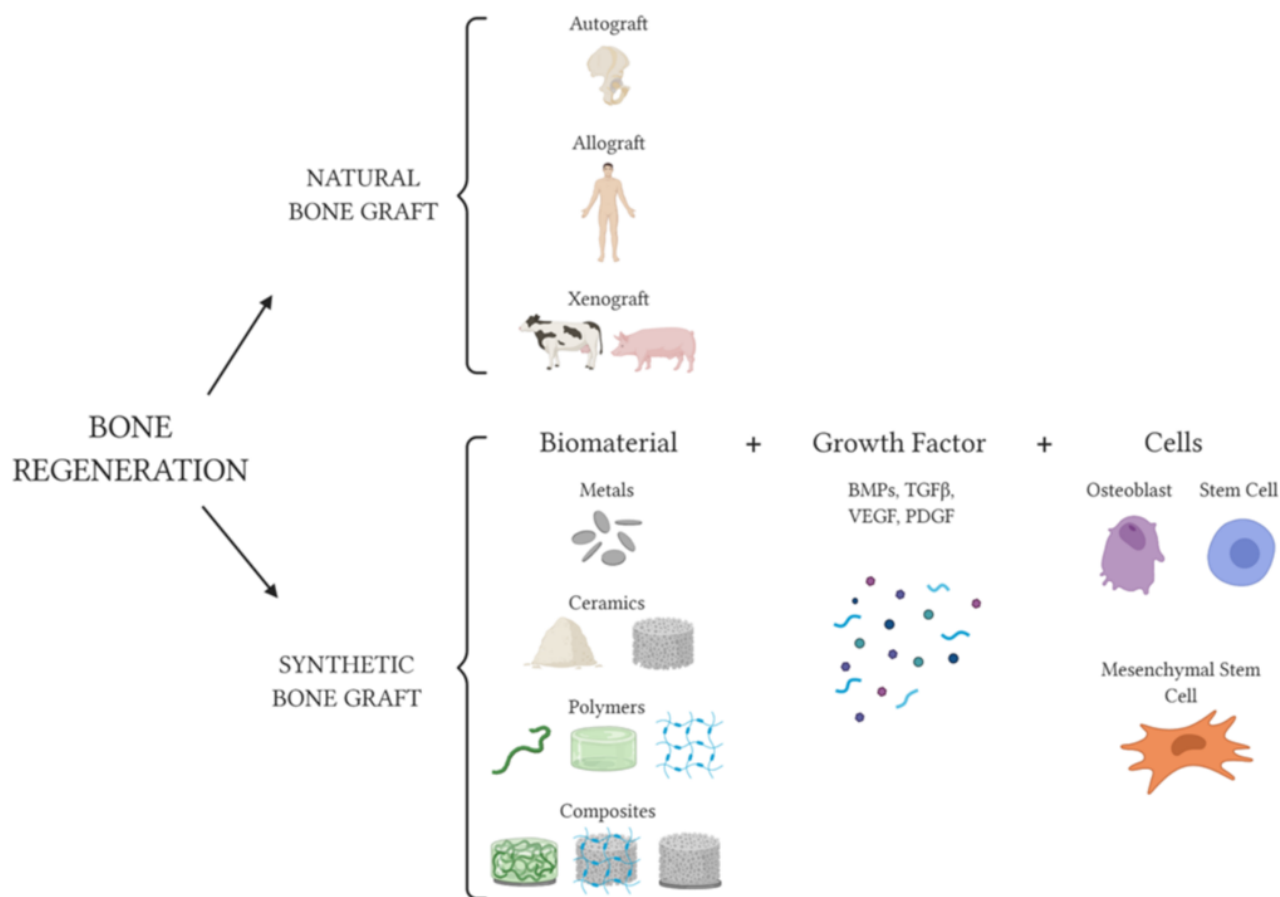


Figure 1.8 Schematic of the bone regeneration treatment options. Natural bone grafts include autografts, allografts and xenografts. New advances are focused on synthetic grafts that comprise biomaterials such as metals, ceramics, polymers, and composites. They may also be cell-seeded or contain growth factors. Adapted from Battafarano et al.⁸⁰.

The use of growth factors in regenerative medicine has become a compelling topic of research after the US Food and Drug Administration (FDA) approved the use of

recombinant human bone morphogenetic protein 2 (rhBMP2) for limited orthopaedic applications in 2002⁸¹. In both 2004 and 2007, the FDA has extended their approval of rhBMP2 as treatment for acute open tibial fractures and bone grafting alternative for alveolar ridge augmentation^{82, 83}. In its early days, the administration of growth factors as therapeutics has focused mostly on bolus injections or systemic intravenous administration. However, these proteins have a short half-life and tend to diffuse away from their target tissues. Previous studies showed that only 0.1% of the dose was present at the target tissue 24 hours after administration⁸⁴. The short half-life of BMP2 and its fast diffusion amount to the reason why supraphysiological concentrations are administered, which in turn can result in undesirable side-effects and considerably increase treatment costs^{85, 86}. For example, the cost of a single treatment can amount to \$3,500-4,900 (~£3,000-4,000)⁸⁷.

The challenges associated with BMP2 treatment arise due to the way the molecules are delivered and the resulting pharmacokinetics. The administration and mobility shortcomings of BMP2 can be addressed by incorporating the growth factor with carriers, such as scaffolds⁸⁸, hydrogels⁸⁹, or microparticles⁹⁰. Turner *et al.* created gelatin microspheres for the entrapment of VEGF and BMP2 with the use of genipin crosslinkers⁹¹. The release of the growth factors was determined by the hydrogel degradation rate using collagenase. In another study, Xiong *et al.* developed a double network hydrogel loaded with magnesium ion and BMP2 to work as a synergistic treatment for bone defects⁹². Although, the release of magnesium ions and BMP2 was controlled by hydrogel degradation, they were able to elicit a degree of cell proliferation and osteogenic differentiation *in vivo* and *in vitro*. Many other studies on growth factor hydrogels focus on the degradation kinetics as a release mechanism, however, with the process of bone healing, the spatio-temporal control is crucial, and this aspect should be investigated further. To provide a more physiological, sustained, and cost-effective

alternative for BMP2 delivery to the defect site, a growth factor could be produced *in-situ*, potentially using a method such as cell-free protein synthesis.

1.2 CELL-FREE PROTEIN SYNTHESIS

1.1.5 PROTEIN SYNTHESIS IN E.COLI

Gene expression is a process of converting genetic information stored in the form of DNA, firstly into RNA, and subsequently translating into a protein. In eukaryotes, transcription occurs within the nucleus and the generated messenger RNA (mRNA) is selectively transported through nuclear pores into the cytoplasm, allowing translation to take place. Whereas, in prokaryotes, the transcription and translation can occur simultaneously due to the lack of a nucleus. The mechanism of transcription is primarily orchestrated by RNA polymerase (RNAP) as shown in Figure 1.9. This enzyme is composed of five subunits $\alpha_2\beta\beta'\omega$ creating the core polymerase⁹³. The responsibility of the core enzyme is the synthesis of RNA from the DNA template by polymerising ribonucleoside 5'-triphosphates (NTPs) in the 5' to 3' direction. A holoenzyme is formed when a σ subunit associates with the core polymerase. It is the σ subunit that recognises the -35 and -10 promoter regions upstream from the gene and allows for sequence specific binding of the RNAP with higher affinity⁹⁴. There are several σ factors that have been identified both in *E.coli* and bacteriophage, and each of them is responsible for transcription initiation of specific set of genes, with sigma factor 70 (σ^{70}) being the most studied activator⁹⁵. Once the binding of RNAP to the respective regions of DNA take place, the enzyme unwinds ~15 bases of DNA around the initiation element, exposing a single-stranded template. The enzyme then begins to join up free NTPs and the σ factor dissociates from the complex, allowing RNAP to travel further along the template always unwinding DNA ahead of it and rewinding it behind. This process of NTP polymerisation continues until RNAP encounters one of the termination strategies⁹⁶. In Rho-dependent termination, a Rho factor attaches to a newly formed mRNA sequence and travels in the 5'-3' direction towards the RNAP-DNA complex. Once a Rho factor reaches it, the complex dissociates,

and mRNA becomes released. In Rho-independent termination, a transcription of a guanine and cytosine-rich inverted repeat results in a formation of a self-complementary hairpin that induces stalling in RNAP, and ultimately its dissociation.

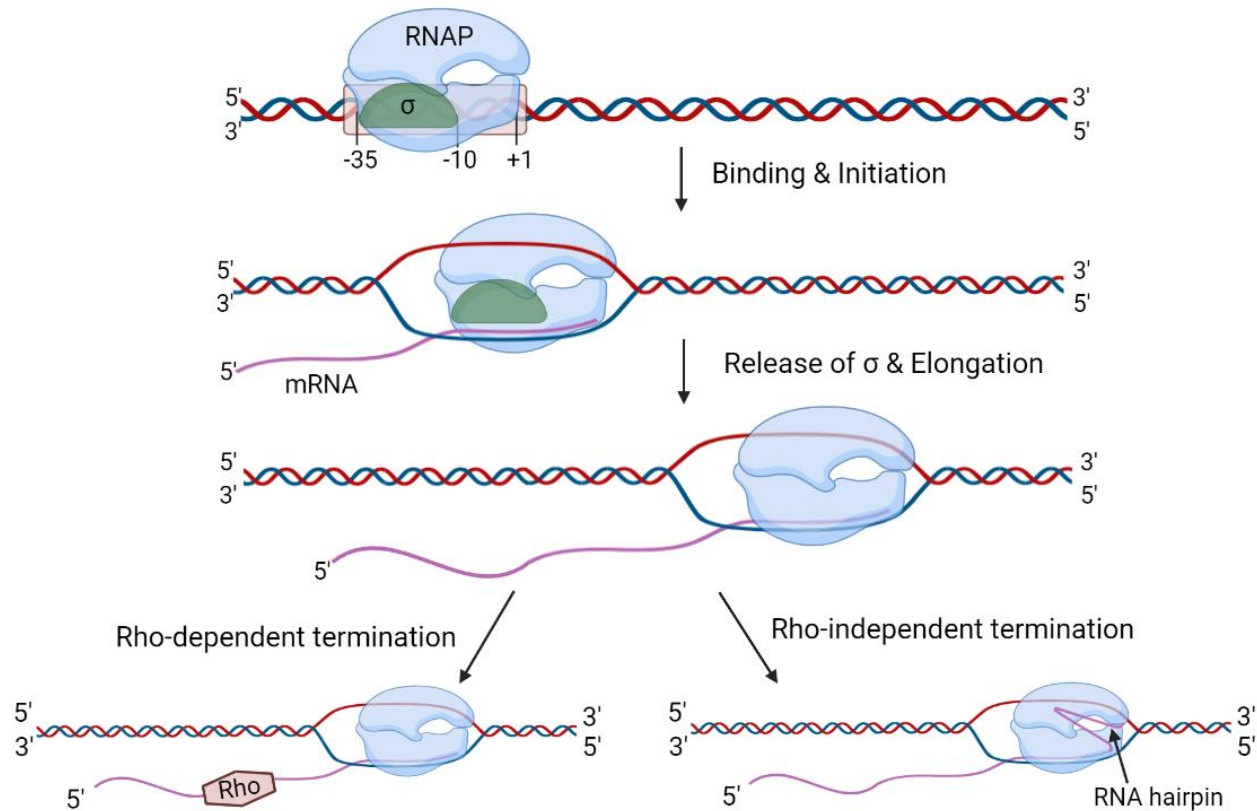


Figure 1.9 Schematic of the prokaryotic transcription by RNA polymerase. The sigma factor allows for sequence specific binding of core RNAP at the promoter. The holoenzyme unwinds DNA and begins transcription by joining NTPs. Sigma factor is released and the core RNAP travels along the DNA elongating the transcript. Termination occurs by Rho factor or by formation of an RNA hairpin. Figure created in biorender.com.

The resulting mRNA transcript is then translated from RNA into a polypeptide chain formed from amino acids. The process of translation can also be divided into initiation, elongation, and termination. However, the whole operation is much more complex than transcription, involving multiple components and a larger pool of substrates, with some of the major components being the mRNA, transfer RNAs (tRNAs), aminoacyl-tRNA synthetase and the ribosome. In prokaryotes, mRNAs not only contain a start codon, usually a 5'-AUG-3', but also a purine-rich ribosome-binding site (RBS), also called the

Shine-Dalgarno sequence, located 5-9 nucleotides upstream of the start codon⁹⁷. It has been found that the sequence at RBS can influence the rate of translation initiation, although it is not the sole factor, with a type of initiation codon and a spacer sequence between RBS and the start codon also being at play⁹⁸. In eukaryotes, mRNA is processed to contain a 5'-end m⁷G-cap and a 3'-end poly(A) tail, both stabilising the transcript and regulating translation efficiency⁹⁹. A ribonucleoprotein, the ribosome, is composed of two parts: a large 50S (60S in eukaryotes) subunit and a small 30S (40S in eukaryotes) subunit, assembling into a 70S ribosome that is central in the process of translation¹⁰⁰. The two subunits associate together at translation initiation and remain bound until termination of synthesis occurs.

The role of the ribosome is to catalyse the formation of the peptide bond between aminoacyl-tRNA and peptidyl-tRNA. In prokaryotes, certain initiation factors bind to the small ribosomal subunit and together attach to the RBS on mRNA (Figure 1.10). The initiator tRNA, fMet-tRNA^{fMet}, then takes a position aligning with the complementary start codon and finally, to complete the initiation complex, the 50S subunit assembles at the top by expelling the initiation factors¹⁰¹. The elongation then proceeds by the arrival of aminoacyl-tRNA with an elongation factor and GTP. When the GTP is hydrolysed, the elongation factor undergoes an exchange cycle¹⁰². A peptide bond between the amino acids is formed by the activity of the peptidyl transferase and upon energy consumption the ribosome moves along the mRNA by one codon and the now 'uncharged' tRNA is removed¹⁰³. Termination of translation occurs when a release factor interacts with a stop codon (UAA, UAG, UGA and UAA) leading to a water molecule addition to the polypeptide chain by the peptidyl transferase on the ribosome. The assembled complex is dissociated with the help of ribosome recycling factors¹⁰⁴.

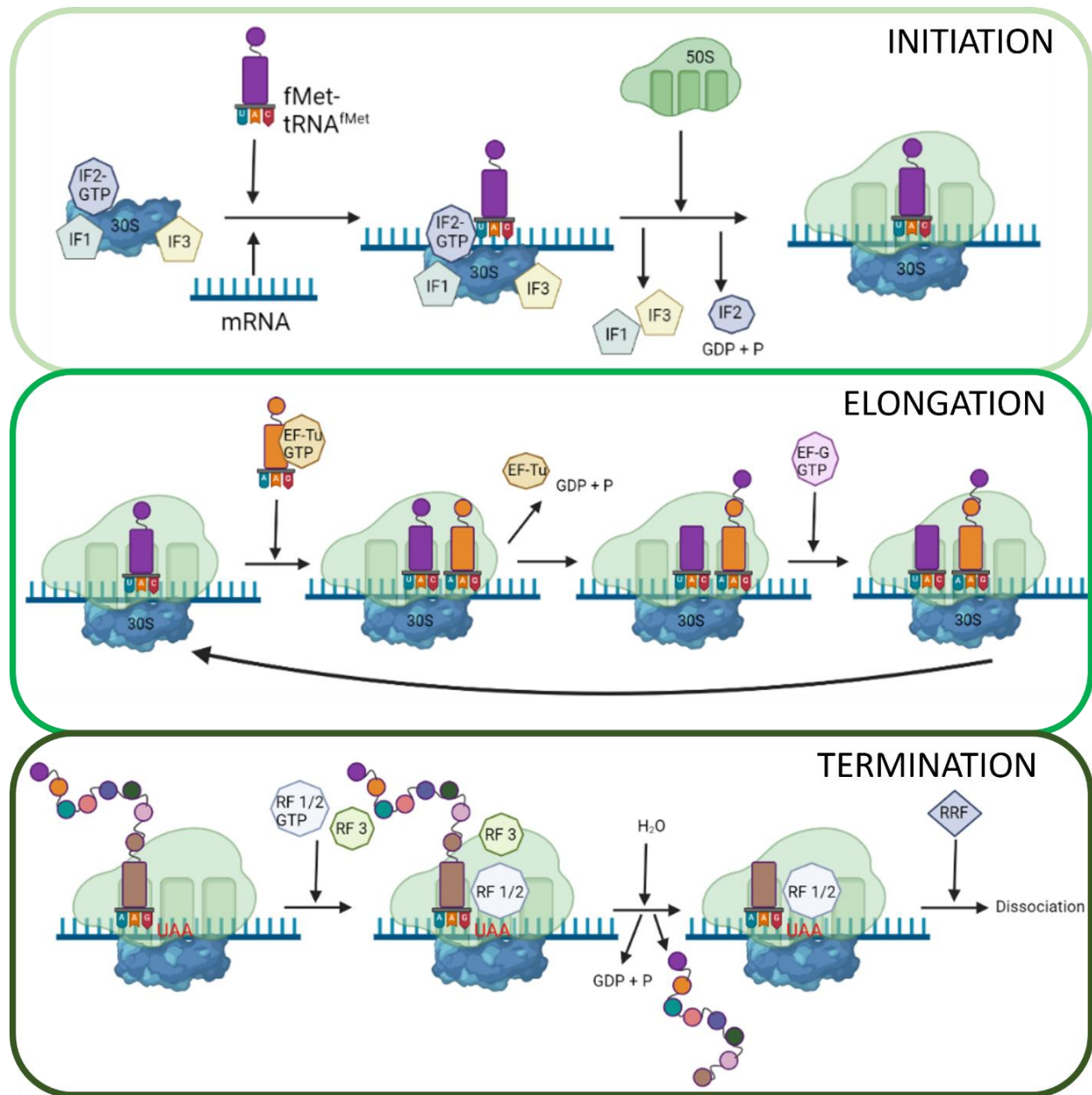


Figure 1.10 Schematic of the process of prokaryotic translation. In initiation, ribosome subunits come together allowing for the initiator tRNA to recognise the start codon. In elongation, new tRNAs complementary base pairs with codons and the amino acids they carry bind to the existing chain *via* a peptide bond. Termination is reached when the stop codon is recognised by a release factor. Hydrolysis of an ester bond leads to the polypeptide chain release whilst a ribosome recycling factor dissociates the ribosome complex from mRNA. Figure created in biorender.com.

The newly synthesised polypeptide chain will then undergo the process of physical folding to form a biologically functional 3D structure. Despite considerable number of studies

conducted on protein fold predictions, it still presents as a major challenge, without a unifying mechanism¹⁰⁵. Nonetheless, one theory hypothesises that the folding of the protein follows a funnel-shaped energy landscape with the native state corresponding to the free energy minimum¹⁰⁶. Generally, protein structure has been organised into four levels: primary structure, which is the sequence of amino acids; secondary structure relating to the spatial arrangement of the amino acids, such as the α -helix and β -strand; tertiary structure corresponding to the 3D structure of the polypeptide chain; and the quaternary structure, which is relating the spatial arrangement of subunits within a multi-subunit protein¹⁰⁷.

1.1.6 HISTORY AND COMPONENTS OF THE CELL-FREE PROTEIN SYNTHESIS

Cell-free protein synthesis (CFPS), also referred to as cell-free expression system (CFES), utilise *in vitro* transcription-translation methodologies and date back to the 1950s, when they were applied for understanding basic concepts in molecular biology. The earliest studies conducted to elucidate the way in which amino acids were converted to proteins were performed with rat liver cells extracts¹⁰⁸. Some of the others included bacteria¹⁰⁹, wheat germ¹¹⁰ and human reticulocytes¹¹¹. The early extracts contained endogenous mRNA resulting in the generation of native proteins. It was in 1961, using an S30 cell extract when Nirenberg and Matthaei were successful at deciphering that a specific codon was generating a specific amino acid¹¹². Thanks to this breakthrough, the remaining possible codon combinations, and the amino acids they encoded were unravelled over the years. The S30 extract that was used by Nirenberg and Matthaei was developed from cell lysates that were centrifuged at 30,000 g to generate the used supernatant, and this extract naming convention remains. At this point, Nirenberg and Matthaei introduced several modifications to the extract to encourage expression from exogenous mRNA, including addition of components for ATP regeneration. In 1983, Chen and Zubay coupled transcription with translation by adding DNA, rather than exogenous mRNA into the S30

extract and utilising the function of *E.coli* RNA polymerase¹¹³. Since then, CFPS started to become considered as an *in vitro* protein synthesis method rather than solely an analytical tool. The basic components of a cell-free protein synthesis reaction can be classified into input DNA, the cellular extract and a buffer mix supplemented with energy and amino acids (Figure 1.11). The crude cell extract contains ingredients such as initiation, elongation and release factors, ribosomes, RNA polymerase, tRNAs, and NTPs amongst other transcription-translation (TX-TL) machinery, which initiate protein synthesis when mixed with DNA template, amino acids and an energy solution.

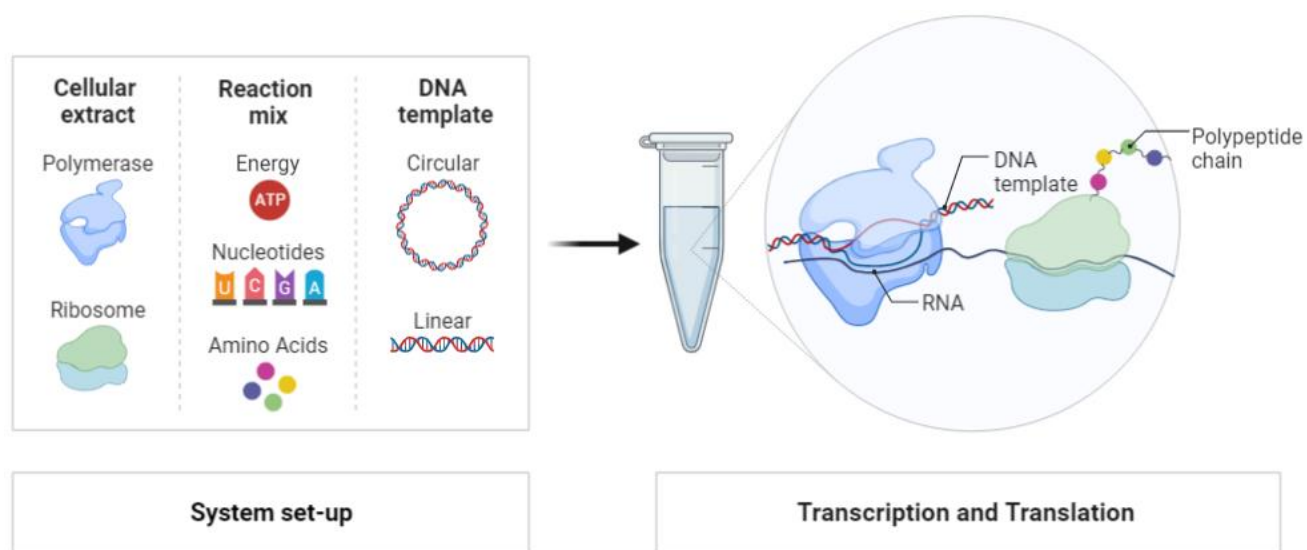


Figure 1.11 A basic coupled transcription-translation reaction set-up. Protein synthesis can be achieved by mixing cellular extract, input DNA and a buffer reaction mix. Figure generated from biorender.com.

A cell-free extract protocol can be broadly divided into initial cell growth, lysis, and multi-step clarification. Improvements in most areas of the cell-free extract protocol have been attempted. Some of the alterations in the improved protocols included the use of stronger phage promoter or RNA polymerases such as T7¹¹⁴ or T3¹¹⁵, different strains of cells and various lysis methods such as sonication¹¹⁶, French press¹¹⁷, or bead-beating¹¹⁸. The alterations also extend to the components that were found to play a role in protein synthesis and some of these components are summarised in Table 1.1. Various components

can be used to optimise the environment of the cell-free reaction, all with the goal of increasing the protein production yield or reducing the running costs.

Component	Role in CFPS
Amino Acids	Substrates for translation
ATP	Energy source and a substrate for transcription
cAMP	Increased yields
Coenzyme A	Regeneration of ATP from pyruvate
NAD	
Creatine phosphate	Energy regeneration, phosphorylation of nucleoside monophosphates
Phosphoenolpyruvate	
Pyruvate	
3-Phosphoglyceric acid	
DTT	Stabilises polymerases
Folinic acid	Formation of initiator formyl-methionine
GTP, CTP, UTP	Substrates for transcription
HEPES	Buffer
Maltose	Recycling of inorganic phosphate
Mg ²⁺ , K ⁺ , NH ₄ ⁺ , glutamate, acetate	Ions
Nucleoside monophosphates	Cheaper source of substrates for transcription
Oxalate	Inhibitor of phosphoenolpyruvate synthetase
Polyethylene glycol	Viscosity modifier, stabilises mRNA, crowding effect
Putrescine	Stabilisation of nucleic acids, stimulate polymerase activity
Spermidine	
tRNA	Supply of Amino Acids

Table 1.1 A list of some components that are added to certain cell extracts, and their roles in increasing the yield of the synthesised proteins. Table has been adapted from Rolf *et al.*¹¹⁹.

1.1.7 ADVANTAGES, LIMITATIONS AND APPLICATIONS

In cell-free protein synthesis, the removal of the native genomic DNA means that the necessity of cell viability is eliminated, and the genetic regulation of housekeeping genes is abolished, leading to more flexibility in terms of engineering¹²⁰. The absence of the cell membrane and organelles provides openness to the experimental environment, allowing whole sample monitoring as well as easy manipulation of substrates and products¹²¹. This is a particularly useful approach for obtaining proteins that are toxic to the host cell¹²², difficult-to-express proteins¹²³, as well as membrane proteins when lipids are supplemented¹²⁴. When recombinant toxic proteins undergo *in vivo* synthesis, it usually results in the cell's death or inhibition of production and therefore, very minimal yields¹²⁵. Additionally, *in vivo* expression of membrane proteins can reduce cell growth due to disordered cell metabolism, cause cytotoxicity or simply lead to formation of protein aggregates¹²⁶. However, these limitations are altogether diminished in a cell-free protein synthesis system. When membrane proteins are expressed *via* the cell-free method, the topology or size of the protein of interest is not under restriction. Furthermore, the proteins can be translated directly into detergent micelles, unconstrained by the translocation complexes of the cell for membrane protein insertion into a membrane¹²⁷. Furthermore, CFPS allows for the transcription of linear or circular DNA and mRNA without the need of transformation or transfection of cells, leading to rapid and high-throughput protein production. The advantages of this flexible and customisable expression are especially appealing for on-demand production of a certain dose of proteins, such as pharmaceuticals or in the case of this thesis, growth factors.

Despite the advantageous CFPS over traditional *in vivo* protein expression, there are still certain limitations associated with its use. These restrictions include poor scalability, especially with eukaryotic systems. Moreover, some posttranslational modifications are more difficult to obtain in a cell-free system, notably glycosylation. The production cost is also higher when the added components for TX-TL are taken into consideration¹²⁸.

Contradictory to a living cell, the energy sources in CFPS become used up and are not continuously regenerated. Therefore, one of the areas of CFPS that is highly researched and frequently improved upon is the energy regeneration. For example, it has been found that using phosphoenolpyruvate as an energy source can be inhibitory to CFPS due to rapid production of phosphates which sequesters magnesium ions, only providing a short burst of energy¹²⁹. This can be challenged by adding an inhibitor of phosphoenolpyruvate synthetase, called oxalate¹³⁰, and by recycling the phosphate by phosphorylation of added maltose¹³¹. Underwood *et al.* found that a major limiting step in *E.coli* CFPS is the translation phase, with almost 200-fold slower rate of synthesis when compared to the *in vivo* counterpart¹³². By probing with polysome profile analysis combined with protein quantification performed by Underwood *et al.*, the elongation rate of ribosomes was identified to be limiting the translation in the system. To dissipate this discrepancy between *in vitro* and *in vivo* translation rates, Nieß *et al.* proposed additional supply of elongation factor Tu (EFTu) and tRNAs¹³³.

Currently, *E.coli* based CFPS remains the most common system for cell-free expression, but since 1960s technological advances have opened up numerous applications (Figure 1.12). A number of protocols have been published for effective in-house production of cell-free extract, yet with the rise in CFPS interest, commercial kits are available, enabling even easier access to cell-free biology. It is also possible to initiate TX-TL from reconstituted purified components which are marketed as ‘PURExpress’ (PURE standing for protein synthesis using recombinant elements)¹³⁴. The main advantage of using highly purified components rather than a cell extract, is that the system is free from nucleases or proteases. The components that are found in the PURE system have been engineered with a polyhistidine tag. This feature has been utilised to create a protein-expressing cell-mimic by Zhou *et al.*¹³⁵. Here, the PURExpress His-tagged TX-TL components were attached to the Ni²⁺-NTA polymer backbone of a hydrogel particle, enabling continuous uptake of nutrients into and synthesised products out of the centre. It was reported that

the self-regulating artificial cell with an *in vitro* genetic oscillator was capable of stable protein expression up to 11 days. The authors envision the CSPS cell-mimetic particles becoming useful for applications such as biosensors or drug delivery vehicles.

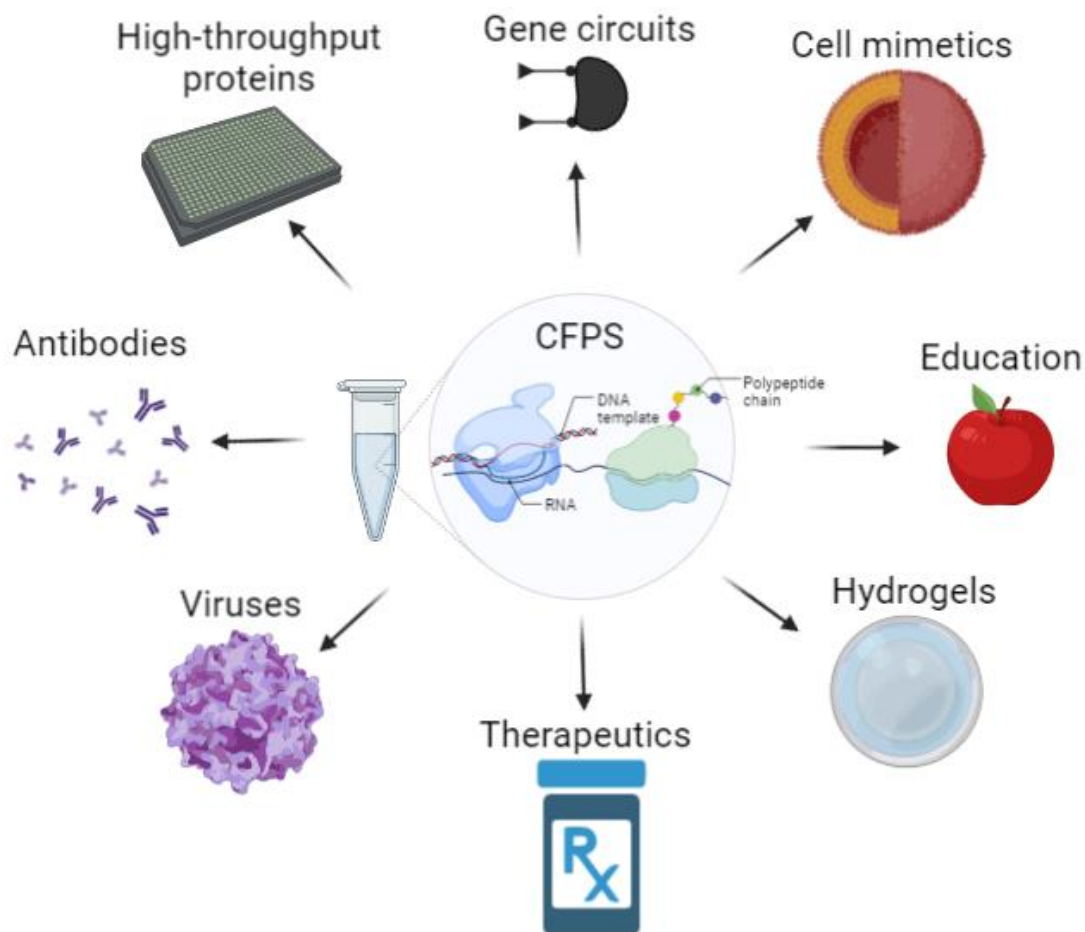


Figure 1.12 A repertoire of applications made possible from cell-free protein synthesis. Some of the applications include high-throughput production, generation of antibodies, viruses or virus-like particles, therapeutics, study of gene circuits, implementation into hydrogels or artificial cells, and serving as an educational tool. Figure created in biorender.com.

CFES applications are immense, beginning with helping to elucidate the importance of localised concentration and protection of biomolecules in early life evolution, as illustrated in 2013 by Yang *et al.*¹³⁶, to controlling metabolic pathways for higher conversion efficiencies, of up to 67-fold, as shown by Rollin *et al.* in 2015¹³⁷. The Noireaux research group has become a modern-day pioneer in regard to gene circuit prototyping within cell-

free expression systems. Over the years, Noireaux *et al.* implemented activation and repression cascades in wheat germ extract as well as both T7-based and native *E.coli* extracts, scaling up to switches, logic gates, toeholds amongst others¹³⁸. The generated ‘toolboxes’ prove the possibility of using CFES as a ‘biomolecular breadboard’, expanding the engineering possibilities of synthetic biology.

CFES are mostly solution based, however, new technology advances granted exploration into other mediums which play an important role in protein yield and functionality. In 2014, Pardee *et al.* developed a paper-based CFPS, where the components of TX-TL were freeze-dried onto paper and synthetic gene networks were programmed to result in a colorimetric output following rehydration¹³⁹. This was particularly revolutionary, enabling easier distribution of such technologies as *in vitro* diagnostics or a safe educational tool in schools. On the other hand, Park *et al.* created a ‘P-gel’ where a hydrogel composed of X-branched DNA crosslinkers and gene monomers were embedded in polydimethylsiloxane (PDMS), which was incubated in a TX-TL solution to produce a high-yielding Rluc protein¹⁴⁰. It was found that this format was providing DNA protection and the physical crosslinking of the genes into the hydrogel led to heightened protein synthesis, compared to a solution-phase CFPS. A similar finding was established by Whitfield *et al.*, where cell-free protein synthesis was investigated in a range of different hydrogel materials¹⁴¹. CFES was successfully incorporated into polysaccharide, proteinaceous, covalently cross-linked, and micellar microscale hydrogels, with agar and agarose yielding the highest protein synthesis, surpassing that of the solution-phase CFPS. Whitfield *et al.* was also able to demonstrate the ability of the gels to act as a molecular crowding agent, thereby having a positive impact on CFPS efficiency.

It is becoming clear that temporal and spatial delivery of growth factors such as BMP2 is key for advanced *in vivo* bone tissue regenerative therapies. CFES provides a possible solution to this problem, as the cell-free method for protein production can be highly controlled.

1.3 PROJECT DESIGN AND AIMS

As outlined in the literature discussion of this chapter, the use of growth factors as therapeutics for regenerative medicine purposes, is effective at inducing new bone formation. However, due to the challenges associated with administering the growth factor, a safer, more precise, and controllable method needs to be constructed. In this project, the overall aim is to provide groundwork on developing an injectable gel encasing a cell-free expression system capable of on-demand production of a growth factor *in-situ*. This methodology would combat the current therapeutic problems of rapid diffusion and supra-physiological doses used. Vascular endothelial growth factor (VEGF) and bone morphogenetic protein 2 (BMP2) are of particular interest due to their ability to stimulate the generation of a vascular network needed for the oxygen and nutrient supply, and the differentiation of mesenchymal stem cells into osteoblasts and their proliferation during bone regeneration, respectively. This process is tightly controlled; and therefore, in an ideal system, the growth factor expression should be modulated by a switch once in place at the fracture site, in order to replicate the natural course of bone healing. It is speculated that the resulting growth-factor producing hydrogel will enhance bone repair by stimulating human mesenchymal stem cells into osteogenic differentiation. Commercialisation of this treatment would require multiple years of research, therefore, this project focuses on some of the initial steps towards the ultimate goal. The work presented here aims to show the:

- Creation of a growth factor and reporter protein fusion compatible with cell-free protein expression.
- Delineation of the cell-free expression capabilities.
- Testing of the system's differentiation potential with human mesenchymal stem cells.
- Incorporation of the cell-free expression system into various gels and their efficiency comparisons

1.4 THESIS STRUCTURE

The work presented in this thesis consists of four results chapters, each opening with a short introduction to the requisite background information, followed by materials and methods used therein. Results and detailed discussion of the data are then presented, followed by conclusions with suggestions for future work. Consecutive to the results chapters, is a general conclusions chapter containing a summary of the main results and a brief discussion on the next steps towards the end goal. In the first chapter of results, Chapter 2, the design and cloning of several cell-free protein synthesis compatible plasmids of a reporter protein and growth factor are detailed. In Chapter 3, those plasmids are compared for expression yields using cell-free protein synthesis. In the same chapter, incorporation of cell-free reactions into three different gels is briefly investigated. In Chapter 4, the highest-yielding reporter protein and growth factor fusion plasmid, deGFP-BMP2, is expressed in *E.coli* and characterised for its structural and biophysical properties. The final results chapter, Chapter 5, explores the interactions and differentiation potential of deGFP-BMP2 on human mesenchymal stem cells.

1.4 BIBLIOGRAPHY

1. Panteli, M., Pountos, I., Jones, E. & Giannoudis, P. v. Biological and molecular profile of fracture non-union tissue: current insights. *Journal of Cellular and Molecular Medicine* 19, 713 (2015).
2. Hutchison, C., Pilote, M. & Roy, S. The axolotl limb: A model for bone development, regeneration and fracture healing. *Bone* 40, 45–56 (2007).
3. Rundle, C. H. *et al.* Microarray analysis of gene expression during the inflammation and endochondral bone formation stages of rat femur fracture repair. *Bone* 38, 521–529 (2006).
4. Gerstenfeld, L. C., Cullinane, D. M., Barnes, G. L., Graves, D. T. & Einhorn, T. A. Fracture healing as a post-natal developmental process: Molecular, spatial, and temporal aspects of its regulation. *Journal of Cellular Biochemistry* 88, 873–884 (2003).
5. Kolar, P. *et al.* The Early Fracture Hematoma and Its Potential Role in Fracture Healing. *Tissue Engineering Part B: Reviews* 16, 427–434 (2010).
6. Loi, F. *et al.* Inflammation, fracture and bone repair. *Bone* 86, 119–130 (2016).
7. Einhorn, T. A. & Gerstenfeld, L. C. Fracture healing: mechanisms and interventions. *Nature Reviews Rheumatology* 11, 45–54 (2015).
8. Phillips, A. M. Overview of the fracture healing cascade. *Injury* 36, S5–S7 (2005).
9. Sandberg, O. Distal radial fractures heal by direct woven bone formation. *Acta Orthopaedica* 84, 297–300 (2013).
10. Carano, R. & Filvaroff, E. Angiogenesis and bone repair. *Drug Discovery Today* 8, 980–989 (2003).

11. Marsell, R. & Einhorn, T. A. The biology of fracture healing. *Injury* 42, 551–555 (2011).
12. Mountziaris, P. M. & Mikos, A. G. Modulation of the inflammatory response for enhanced bone tissue regeneration. *Tissue Engineering - Part B: Reviews* vol. 14 179–186 Preprint at <https://doi.org/10.1089/ten.teb.2008.0038> (2008).
13. Yu-Yahiro, J. A. *et al.* Serum and Urine Markers of Bone Metabolism During the Year After Hip Fracture. *J Am Geriatr Soc* 49, 877–883 (2001).
14. Ding, D.-C., Shyu, W.-C. & Lin, S.-Z. Mesenchymal Stem Cells. *Cell Transplantation* 20, 5–14 (2011).
15. Pittenger, M. F. *et al.* Multilineage Potential of Adult Human Mesenchymal Stem Cells. *Science (1979)* 284, 143–147 (1999).
16. Huang, G. T. J., Gronthos, S. & Shi, S. Mesenchymal stem cells derived from dental tissues vs. those from other sources: Their biology and role in Regenerative Medicine. *Journal of Dental Research* 88, 792–806 (2009).
17. Schwab, K. E., Hutchinson, P. & Gargett, C. E. Identification of surface markers for prospective isolation of human endometrial stromal colony-forming cells. *Human Reproduction* 23, 934–943 (2008).
18. Fukuchi, Y. *et al.* Human Placenta-Derived Cells Have Mesenchymal Stem/Progenitor Cell Potential. *Stem Cells* 22, 649–658 (2004).
19. Baksh, D., Yao, R. & Tuan, R. S. Comparison of Proliferative and Multilineage Differentiation Potential of Human Mesenchymal Stem Cells Derived from Umbilical Cord and Bone Marrow. *Stem Cells* 25, 1384–1392 (2007).

20. Ermolaeva, M., Neri, F., Ori, A. & Lenhard Rudolph, K. Cellular and epigenetic drivers of stem cell ageing. *Nature Reviews Molecular Cell Biology* 19, 594–610 (2018).
21. Vega, A. *et al.* Treatment of knee osteoarthritis with allogeneic bone marrow mesenchymal stem cells: A randomized controlled trial. *Transplantation* 99, 1681–1690 (2015).
22. Dawn, B., Abdel-Latif, A., Sanganalmath, S. K., Flaherty, M. P. & Zuba-surma, E. K. Cardiac Repair with Adult Bone Marrow-Derived Cells: The Clinical Evidence. *Antioxidants & Redox Signaling* 11, 1865–1882 (2009).
23. Liao, Y., Zhang, X. L., Li, L., Shen, F. M. & Zhong, M. K. Stem cell therapy for bone repair: a systematic review and meta-analysis of preclinical studies with large animal models. *British Journal of Clinical Pharmacology* 78, 718–726 (2014).
24. Pankajakshan, D. & Agrawal, D. K. Mesenchymal Stem Cell Paracrine Factors in Vascular Repair and Regeneration. *Journal of Biomedical Technology and Research* 1, (2014).
25. Park, K. S., Bandeira, E., Shelke, G. v., Lässer, C. & Lötvall, J. Enhancement of therapeutic potential of mesenchymal stem cell-derived extracellular vesicles. *Stem Cell Research and Therapy* vol. 10 Preprint at <https://doi.org/10.1186/s13287-019-1398-3> (2019).
26. von Bahr, L. *et al.* Analysis of Tissues Following Mesenchymal Stromal Cell Therapy in Humans Indicates Limited Long-Term Engraftment and No Ectopic Tissue Formation. *Stem Cells* 30, 1575–1578 (2012).
27. Zhao, J., Zhang, N., Prestwich, G. D. & Wen, X. Recruitment of endogenous stem cells for tissue repair. *Macromolecular Bioscience* 8, 836–842 (2008).

28. Lv, F. J., Tuan, R. S., Cheung, K. M. C. & Leung, V. Y. L. The Surface Markers and Identity of Human Mesenchymal Stem Cells. *Stem Cells* 32, 1408–1419 (2014).
29. Amarasekara, D. S., Kim, S. & Rho, J. Regulation of Osteoblast Differentiation by Cytokine Networks. *International Journal of Molecular Sciences* 22, (2021).
30. Greenhalgh, D. G. The Role of Growth Factors in Wound Healing. *The Journal of Trauma: Injury, Infection, and Critical Care* 41, 159–167 (1996).
31. Hojo, H., Ohba, S. & Chung, U. il. Signaling pathways regulating the specification and differentiation of the osteoblast lineage. *Regenerative Therapy* vol. 1 57–62 Preprint at <https://doi.org/10.1016/j.reth.2014.10.002> (2015).
32. Hollinger, J. O. *et al.* Therapeutic Opportunities for Bone Grafting. in *Principles of Regenerative Medicine* 1164–1175 (2008).
33. Khalil, N. TGF- β : from latent to active. *Microbes and Infection* 1, 1255–1263 (1999).
34. Sartori, R. *et al.* BMP signaling controls muscle mass. *Nature Genetics* 45, 1309–1318 (2013).
35. Gustafson, B. *et al.* BMP4 and BMP Antagonists Regulate Human White and Beige Adipogenesis. *Diabetes* 64, 1670–1681 (2015).
36. Andl, T. *et al.* Epithelial Bmpr1a regulates differentiation and proliferation in postnatal hair follicles and is essential for tooth development. *Development* 131, 2257–2268 (2004).
37. Nishinakamura, R. & Sakaguchi, M. BMP signaling and its modifiers in kidney development. *Pediatric Nephrology* 29, 681–686 (2014).
38. Cai, J., Pardali, E., Sánchez-Duffhues, G. & ten Dijke, P. BMP signaling in vascular diseases. *FEBS Letters* 586, 1993–2002 (2012).

39. Katagiri, T. & Watabe, T. Bone Morphogenetic Proteins. *Cold Spring Harbour Perspectives in Biology* 8, (2016).
40. Urist, M. R. & Strates, B. S. Bone Morphogenetic Protein. *Journal of Dental Research* 50, 1392–1406 (1971).
41. Wang, E. A. *et al.* Purification and characterization of other distinct bone-inducing factors. *Proc Natl Acad Sci U S A* 85, 9484–9488 (1988).
42. Wozney, J. M. The bone morphogenetic protein family and osteogenesis. *Molecular Reproduction and Development* 32, 160–167 (1992).
43. Sykaras, N. & Opperman, L. A. Bone morphogenetic proteins (BMPs): how do they function and what can they offer the clinician? *Journal of Oral Science* vol. 45 (2003).
44. Rosenzweig, B. L. *et al.* Cloning and characterization of a human type II receptor for bone morphogenetic proteins. *Proceedings of the National Academy of Sciences* 92, 7632–7636 (1995).
45. Chen, G., Deng, C. & Li, Y. P. TGF- β and BMP signaling in osteoblast differentiation and bone formation. *International Journal of Biological Sciences* 8, 272–288 (2012).
46. Wang, R. N. *et al.* Bone Morphogenetic Protein (BMP) signaling in development and human diseases. *Genes and Diseases* 1, 87–105 (2014).
47. Scheufler, C., Sebald, W. & Hülsmeier, M. Crystal structure of human bone morphogenetic protein-2 at 2.7 Å resolution. *Journal of Molecular Biology* 287, 103–115 (1999).
48. Mueller, T. D. & Nickel, J. Promiscuity and specificity in BMP receptor activation. *FEBS Letters* 586, 1846–1859 (2012).

49. Rosen, V. & Thies, R. S. The BMP proteins in bone formation and repair. *Trends in Genetics* 8, 97–102 (1992).
50. Zhang, H. & Bradley, A. Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Development* 122, 2977–2986 (1996).
51. McBride, S. H. *et al.* Long Bone Structure and Strength Depend on BMP2 from Osteoblasts and Osteocytes, but Not Vascular Endothelial Cells. *PLOS ONE* 9, (2014).
52. Bandyopadhyay, A. *et al.* Genetic Analysis of the Roles of BMP2, BMP4, and BMP7 in Limb Patterning and Skeletogenesis. *PLOS Genetics* 2, (2006).
53. Shu, B. *et al.* BMP2, but not BMP4, is crucial for chondrocyte proliferation and maturation during endochondral bone development. *Journal of Cell Science* 124, 3428–3440 (2011).
54. Rountree, R. B. *et al.* BMP Receptor Signaling Is Required for Postnatal Maintenance of Articular Cartilage. *PLOS Biology* 2, (2004).
55. Sözen, T., Özişik, L. & Başaran, N. Ç. An overview and management of osteoporosis. *Eur J Rheumatol.* 4, 46–56 (2017).
56. Salari, N. *et al.* The global prevalence of osteoporosis in the world: a comprehensive systematic review and meta-analysis. *J Orthop Surg Res* 16, (2021).
57. Melton, L. J., Chrischilles, E. A., Cooper, C., Lane, A. W. & Riggs, B. L. Perspective how many women have osteoporosis? *Journal of Bone and Mineral Research* 7, 1005–1010 (1992).
58. Reginster, J. Y. & Burlet, N. Osteoporosis: A still increasing prevalence. *Bone* 38, 4–9 (2006).

59. Sánchez-Duffhues, G., Hiepen, C., Knaus, P. & ten Dijke, P. Bone morphogenetic protein signaling in bone homeostasis. *Bone* vol. 80 43–59 Preprint at <https://doi.org/10.1016/j.bone.2015.05.025> (2015).
60. Klibanski, A. *et al.* Osteoporosis Prevention, Diagnosis, and Therapy. *JAMA* 285, 785–795 (2001).
61. Garrett, I. R. Anabolic Agents and the Bone Morphogenetic Protein Pathway. *Current Topics in Developmental Biology* vol. 78 127–171 Preprint at [https://doi.org/10.1016/S0070-2153\(06\)78004-8](https://doi.org/10.1016/S0070-2153(06)78004-8) (2007).
62. Tsuji, K. *et al.* BMP2 activity, although dispensable for bone formation, is required for the initiation of fracture healing. *Nature Genetics* 2006 38:12 38, 1424–1429 (2006).
63. Bandyopadhyay, A., Yadav, P. S. & Prashar, P. BMP signaling in development and diseases: A pharmacological perspective. *Biochemical Pharmacology* vol. 85 857–864 Preprint at <https://doi.org/10.1016/j.bcp.2013.01.004> (2013).
64. Rauch, F. & Glorieux, F. Osteogenesis Imperfecta. *The Lancet* 363, 1377–1385 (2004).
65. Roodman, G. D. & Windle, J. J. Paget disease of bone. *Journal of Clinical Investigation* 115, 200–208 (2005).
66. Shaw, K. A., Griffith, M. S., Shaw, V. M., Devine, J. G. & Gloystein, D. M. Harvesting Autogenous Cancellous Bone Graft from the Anterior Iliac Crest. *Clin Orthop Relat Res* 8, 300–309 (2018).
67. García-Gareta, E., Coathup, M. J. & Blunn, G. W. Osteoinduction of bone grafting materials for bone repair and regeneration. *Bone* 81, 112–121 (2015).

68. Oryan, A., Alidadi, S., Moshiri, A. & Maffulli, N. Bone regenerative medicine: classic options, novel strategies, and future directions. *Journal of Orthopaedic Surgery and Research* 9, 1–27 (2014).
69. Ehrler, D. M. & Vaccaro, A. R. The Use of Allograft Bone in Lumbar Spine Surgery. *Clinical Orthopaedics and Related Research* 371, 38–45 (2000).
70. Denry, I. & Kuhn, L. T. Design and characterization of calcium phosphate ceramic scaffolds for bone tissue engineering. in *Dental Materials* (2016). doi:10.1016/j.dental.2015.09.008.
71. Fernandes, H. R. *et al.* Bioactive Glasses and Glass-Ceramics for Healthcare Applications in Bone Regeneration and Tissue Engineering. *Materials* 11, (2018).
72. Yashavantha Kumar, C., Nalini, K. B., Menon, J., Patro, D. K. & Banerji, B. H. Calcium sulfate as bone graft substitute in the treatment of osseous bone defects, a prospective study. *Journal of Clinical and Diagnostic Research* 7, 2926–2928 (2013).
73. Barrère, F., A van Blitterswijk, C. & de Groot, K. Bone regeneration: molecular and cellular interactions with calcium phosphate ceramics. *International Journal of Nanomedicine* 1, 317–332 (2006).
74. Pereira, H. F., Cengiz, I. F., Silva, F. S., Reis, R. L. & Oliveira, J. M. Scaffolds and coatings for bone regeneration. *Journal of Materials Science: Materials in Medicine* 31, (2020).
75. Dhandayuthapani, B., Yoshida, Y., Maekawa, T. & Sakthi Kumar, D. Polymeric Scaffolds in Tissue Engineering Application: A Review. *International Journal of Polymer Science* 2011, 1–19 (2011).

76. Akilbekova, D., Shaimerdenova, M., Adilov, S. & Berillo, D. Biocompatible scaffolds based on natural polymers for regenerative medicine. *International Journal of Biological Macromolecules* 114, 324–333 (2018).
77. Kohane, D. S. & Langer, R. Polymeric Biomaterials in Tissue Engineering. *Pediatric Research* 63, (2008).
78. Hospodiuk, M., Dey, M., Sosnoski, D. & Ozbolat, I. T. The bioink: A comprehensive review on bioprintable materials. *Biotechnology Advances* 35, 217–239 (2017).
79. Armstrong, J. P. K., Burke, M., Carter, B. M., Davis, S. A. & Perriman, A. W. 3D Bioprinting Using a Templated Porous Bioink. *Advanced Healthcare Materials* 5, 1724–1730 (2016).
80. Battafarano, G. *et al.* Strategies for Bone Regeneration: From Graft to Tissue Engineering. *International Journal of Molecular Sciences* 22, 1128 (2021).
81. Khan, S. N. & Lane, J. M. The use of recombinant human bone morphogenetic protein-2 (rhBMP-2) in orthopaedic applications. *Expert Opinion on Biological Therapy* 4, 741–748 (2004).
82. Premarket Approval (PMA). *INFUSE BONE GRAFT*
<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P000054>
.
83. Premarket Approval (PMA). *Infuse Bone Graft, Bone Grafting Material, Dental, With Biologic Component*
<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P050053>
.

84. Laham, R. J. *et al.* Intracoronary and intravenous administration of basic fibroblast growth factor: myocardial and tissue distribution. *Drug Metab Dispos* 27, 821–6 (1999).
85. Zara, J. N. *et al.* High doses of bone morphogenetic protein 2 induce structurally abnormal bone and inflammation in vivo. *Tissue Engineering - Part A* 17, 1389–1399 (2011).
86. Alt, V. & Heissel, A. Economic considerations for the use of recombinant human bone morphogenetic protein-2 in open tibial fractures in Europe: the German model. *Current Medical Research and Opinions* 22, 19–22 (2006).
87. Roberts, T. T. & Rosenbaum, A. J. Bone grafts, bone substitutes and orthobiologics. *Organogenesis* 8, 114–124 (2012).
88. El-Ghannam, A., Ning, C. Q. & Mehta, J. Cyclosilicate nanocomposite: A novel resorbable bioactive tissue engineering scaffold for BMP and bone-marrow cell delivery. *Journal of Biomedical Materials Research* 71, 377–390 (2004).
89. Park, S. H. *et al.* BMP2-modified injectable hydrogel for osteogenic differentiation of human periodontal ligament stem cells. *Scientific Reports* 7, 1–15 (2017).
90. Johannes Braun, B., Khalil, A. S., Murphy, W. L. & Menger, M. BMP-2-coated mineral coated microparticles improve bone repair in atrophic non-unions . *European Cells and Materials* 33, 1–12 (2017).
91. Turner, P. A., Thiele, J. S. & Stegemann, J. P. Growth factor sequestration and enzyme-mediated release from genipin-crosslinked gelatin microspheres. *J Biomater Sci Polym Ed* 28, 1826–1846 (2017).
92. Xiong, A. *et al.* The fabrication of a highly efficient hydrogel based on a functionalized double network loaded with magnesium ion and BMP2 for bone

- defect synergistic treatment. *Materials Science and Engineering: C* 128, 112347 (2021).
93. Sutherland, C. & Murakami, K. S. An Introduction to the Structure and Function of the catalytic core enzyme of Escherichia coli RNA polymerase. *EcoSal Plus* 8, (2018).
94. Tripathi, L., Zhang, Y. & Lin, Z. Bacterial sigma factors as targets for engineered or synthetic transcriptional control. *Frontiers in Bioengineering and Biotechnology* vol. 2 Preprint at <https://doi.org/10.3389/fbioe.2014.00033> (2014).
95. Paget, M. S. B. & Helmann, J. D. The $\sigma 70$ family of sigma factors. *Genome Biology* 4, 1–6 (2003).
96. Jun, S.-H., Warner, B. A. & Murakami, K. S. RNA Polymerase Reaction in Bacteria. in *Encyclopedia of Biological Chemistry* 167–172 (2013).
97. Wen, J. der, Kuo, S. T. & Chou, H. H. D. The diversity of Shine-Dalgarno sequences sheds light on the evolution of translation initiation. *RNA Biology* 18, 1489–1500 (2021).
98. Scharff, L. B. *et al.* Shine-Dalgarno Sequences Play an Essential Role in the Translation of Plastid mRNAs in Tobacco. *The Plant Cell* 29, 3085 (2017).
99. Gallie, D. R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development* 5, 2108–2116 (1991).
100. Petrov, A. S. *et al.* History of the ribosome and the origin of translation. *Proc Natl Acad Sci U S A* 112, 15396–15401 (2015).
101. Laursen, B. S., Sørensen, H. P., Mortensen, K. K. & Sperling-Petersen, H. U. Initiation of Protein Synthesis in Bacteria. *Microbiology and Molecular Biology Reviews* 69, 101–123 (2005).

102. Xu, B., Liu, L. & Song, G. Functions and Regulation of Translation Elongation Factors. *Frontiers in Molecular Biosciences* 8, 1357 (2022).
103. Voorhees, R. M. & Ramakrishnan, V. Structural Basis of the Translational Elongation Cycle . *Annual Review of Biochemistry* 82, 203–236 (2013).
104. Kiel, M. C., Kaji, H. & Kaji, A. Ribosome recycling: An essential process of protein synthesis. *Biochemistry and Molecular Biology Education* 35, 40–44 (2007).
105. Dill, K. A. & MacCallum, J. L. The protein-folding problem, 50 years on. *Science* 338, 1042–1046 (2012).
106. Fersht, A. R. Transition-state structure as a unifying basis in protein-folding mechanisms: Contact order, chain topology, stability, and the extended nucleus mechanism. *Proc Natl Acad Sci U S A* 97, 1525–1529 (2000).
107. Sun, P., Foster, C. & Boyington, J. Overview of Protein Structural and Functional Folds. *Current Protocols in Protein Science* 35, 1711–1789 (2004).
108. Zamecnik, P. C., Frantz, I. D., Loftfield, R. B. & Stephenson, M. L. Incorporation in vitro of radioactive carbon from carboxyl-labelled DL-alanine and glycine into proteins of normal and malignant rat livers. *Journal of Biological Chemistry* 175, 299–314 (1948).
109. Gale, D. E. F. & Folkes, J. P. Effect of nucleic acids on protein synthesis and amniocid incorporation in disrupted staphylococcal cells. *Nature* 173, 1223–1227 (1954).
110. And, A. M. & Feeley, J. Ribosome activation and polysome formation in vitro: requirement for ATP. *Proc Natl Acad Sci U S A* 56, 1770 (1966).
111. Bank, A. & Marks, P. A. Protein synthesis in a cell free human reticulocyte system: ribosome function in thalassemia. *Journal of Clinical Investigation* 45, 330 (1966).

112. Nirenberg, M. W. & Matthaei, J. H. The dependence of cell-free protein synthesis in *E. coli* upon naturally occurring or synthetic polyribonucleotides. *Proc Natl Acad Sci U S A* 47, 1588–1602 (1961).
113. Chen, H.-Z. & Zubay, G. Prokaryotic Coupled Transcription-Translation. *Methods in Enzymology* 101, 674–690 (1983).
114. Nevin, D. E. & Pratt, J. M. A coupled in vitro transcription-translation system for the exclusive synthesis of polypeptides expressed from the T7 promoter. *FEBS Letters* 291, 259–263 (1991).
115. Vacca, R. A., Marra, E., Quagliariello, E. & Greco, M. Increase of both transcription and translation activities following separate irradiation of the in vitro system components with He-Ne laser. *Biochemical and Biophysical Research Communications* 203, 991–997 (1994).
116. Kwon, Y. C. & Jewett, M. C. High-throughput preparation methods of crude extract for robust cell-free protein synthesis. *Scientific Reports* 5, 1–8 (2015).
117. Zubay, G. In vitro synthesis of protein in microbial systems. *Annual Review of Genetics* 7, 267–287 (1973).
118. Shrestha, P., Holland, T. M. & Bundy, B. C. Streamlined extract preparation for *Escherichia coli*-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques* 53, 163–174 (2012).
119. Rolf, J., Rosenthal, K. & Lütz, S. Application of Cell-Free Protein Synthesis for Faster Biocatalyst Development. *Catalysts* 9, (2019).
120. Hodgman, C. E. & Jewett, M. C. Cell-free synthetic biology: Thinking outside the cell. *Metabolic Engineering* 14, 261–269 (2012).

121. Forster, A. C. & Church, G. M. Synthetic biology projects in vitro. *Genome Research* 17, 1–6 (2007).
122. Ramm, F. *et al.* Mammalian cell-free protein expression promotes the functional characterization of the tripartite non-hemolytic enterotoxin from *Bacillus cereus*. *Scientific Reports* 2020 10:1 10, 1–12 (2020).
123. Jin, X. & Hong, S. H. Cell-free protein synthesis for producing ‘difficult-to-express’ proteins. *Biochemical Engineering Journal* vol. 138 156–164 Preprint at <https://doi.org/10.1016/j.bej.2018.07.013> (2018).
124. Coutable, A. *et al.* Preparation of tethered-lipid bilayers on gold surfaces for the incorporation of integral membrane proteins synthesized by cell-free expression. *Langmuir* 30, 3132–3141 (2014).
125. Saïda, F., Uzan, M., Odaert, B. & Bontems, F. Expression of Highly Toxic Genes in *E. coli*: Special Strategies and Genetic Tools. *Current Protein and Peptide Science* 7, 47–56 (2006).
126. Kubicek, J., Block, H., Maertens, B., Spriestersbach, A. & Labahn, J. Expression and Purification of Membrane Proteins. in *Methods in Enzymology* vol. 541 117–140 (Academic Press, 2014).
127. Schneider, B., Junge, F., Shirokov, V., Durst, F., Schwarz, D., Dötsch, V., Bernhard, F. Membrane protein expression in cell-free systems. *Methods in Molecular Biology* vol. 601 165-196 (Humana Press, 2010).
128. Zemella, A., Thoring, L., Hoffmeister, C. & Kubick, S. Cell-Free Protein Synthesis: Pros and Cons of Prokaryotic and Eukaryotic Systems. *ChemBioChem* 16, 2420–2431 (2015).
129. Kim, D. M. & Swartz, J. R. Prolonging Cell-Free Protein Synthesis by Selective Reagent Additions. *Biotechnology Progress* 16, 385–390 (2000).

130. Kim, D. M. & Swartz, J. R. Oxalate improves protein synthesis by enhancing ATP supply in a cell-free system derived from *Escherichia coli*. *Biotechnology Letters* 22, 1537–1542 (2000).
131. Caschera, F. & Noireaux, V. Synthesis of 2.3 mg/ml of protein with an all *Escherichia coli* cell-free transcription–translation system. *Biochimie* 99, 162–168 (2014).
132. Underwood, K. A., Swartz, J. R. & Puglisi, J. D. Quantitative polysome analysis identifies limitations in bacterial cell-free protein synthesis. *Biotechnology and Bioengineering* 91, 425–435 (2005).
133. Nieß, A., Failmezger, J., Kuschel, M., Siemann-Herzberg, M. & Takors, R. Experimentally Validated Model Enables Debottlenecking of in Vitro Protein Synthesis and Identifies a Control Shift under in Vivo Conditions. *ACS Synthetic Biology* 6, 1913–1921 (2017).
134. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nature* 409, 751–755 (2001).
135. Zhou, X., Wu, H., Cui, M., Lai, S. N. & Zheng, B. Long-lived protein expression in hydrogel particles: Towards artificial cells. *Chemical Science* 9, 4275–4279 (2018).
136. Yang, D. *et al.* Enhanced transcription and translation in clay hydrogel and implications for early life evolution. *Scientific Reports* 3, 1–6 (2013).
137. Rollin, J. A. *et al.* High-yield hydrogen production from biomass by in vitro metabolic engineering: Mixed sugars coutilization and kinetic modeling. *Proceedings of the National Academy of Sciences* 112, 4964–4969 (2015).

Chapter 1 - Introduction

138. Garamella, J., Marshall, R., Rustad, M. & Noireaux, V. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synthetic Biology* 5, 344–355 (2016).
139. Pardee, K. *et al.* Paper-Based Synthetic Gene Networks. *Cell* 159, 940–954 (2014).
140. Park, N., Um, S. H., Funabashi, H., Xu, J. & Luo, D. A cell-free protein-producing gel. *Nature Materials* 8, 432–437 (2009).
141. Whitfield, C. J. *et al.* Cell-free protein synthesis in hydrogel materials. *Chemical Communications* 56, 7108–7111 (2020).

CHAPTER 2. PLASMID DESIGN

2.1 INTRODUCTION

Protein synthesis from cell-free systems is programmed by the input DNA. Therefore, the focus of this chapter is the pursuit of a plasmid encoding a growth factor, capable of being expressed by an in-house, extract-based, cell-free expression system (CFES). The most common method of monitoring the synthesis from CFES is by implementing a reporter protein, such as GFP or mCherry into the input DNA¹. As the protein of interest, along with the reporter protein are synthesised and become correctly folded, the increase in fluorescence can be observed in real-time, providing a quantitative measure of protein expression. Another element on the plasmid design that can prove useful for any expression, includes fusion tags, such as maltose-binding protein (MBP), glutathione S-transferase (GST), or a string of histidine residues^{2, 3, 4}. These tags allow for affinity purification, for example, using immobilised ion affinity chromatography with nickel resin. Moreover, they can aid the detection of the protein *via* Western blots, using relevant commercial antibodies. It has also been shown that some of these tags can act as solubility enhancers, especially in the case of MBP⁵. If there is a need of removing any of the additional tags from the recombinant protein, a protease cleavage site can be inserted downstream of the sequence, with tobacco etch virus (TEV) protease being a prevalent choice⁶. A critical aspect of plasmid design is the promoter sequence that is recognised by the RNA polymerase (RNAP) with its associated sigma factor (see Chapter 1, Section 1.1.5 for a detailed description). The most recognised is the lac promoter of the lac operon⁷. The expression from this promoter is constitutive when the lac repressor is absent, but when the repressor is bound to the operator, isopropyl β -D-1-thiogalactopyranoside (IPTG) or lactose can fully relieve this repression⁸. A common promoter, P_{T7}, comes from T7 bacteriophage and can result in constitutive expression, but a specific T7 RNA polymerase is needed⁹. pTac is a hybrid promoter of the lac and the tryptophan operons, with tight regulation¹⁰. Another important promoter that is heavily involved in this thesis

is P_{70-mut}. It can be described as a strong, single-point mutated, Lambda phage promoter with two operators, OR2 and OR1, specific to endogenous *E.coli* sigma factor 70 ¹¹.

Here, plasmids with desirable elements, such as specific promoter regions, a reporter protein and purification tags have been obtained from Addgene or from collaborators, and were rationally modified to include the sequence of the growth factors bone morphogenetic protein 2 (BMP2) and vascular endothelial growth factor (VEGF). The first type of plasmid explored was the pCellFree (pCF) series, where the backbone with the desirable elements were designed by Gagoski *et al.*¹². Plasmids pCellFree carry a Species Independent Translation Initiation Sequence (SITS), enabling the protein synthesis in both prokaryotic and eukaryotic in-vitro TX-TL systems. Plasmids pCellFreeG_03 through to G_06 (see Appendix A) have a T7 promoter driving the transcription, an 8xHis-tag for purification purposes, and a PreScission protease cleavage site sequence between the Open Reading Frame (ORF) of interest and the fusion tags. G_03 vector has EGFP upstream of the ORF, whereas in G_04, the reporter is located downstream. The layout in pCellFreeG_05 and G_06 is respective, although the reporter protein is the mCherry. The ORF in pCF is flanked by attachment (att) sites which are recombination hotspots when in contact with corresponding att sites and an enzyme mix. This implies pCF plasmids are Gateway-cloning compatible. This cloning method was chosen to insert BMP2 sequence into the EGFP containing vectors, and VEGF sequence into the mCherry containing vectors of the pCF series.

Another EGFP vector which required further modification, was an Aquaporin Z- EGFP fusion under a pTar promoter that requires *E.coli* sigma factor 28 (σ 28) for recognition of the RNA polymerase ¹³. This signifies that when this plasmid is applied to the endogenous *E.coli* CFES, the transcription will only take place in the presence of sigma factor 28 sourced from another plasmid, pBEST-p15a-OR2-OR1-Pr-UTR1-sigma28-T500 ¹¹. The additional control over the regulation of the expression plasmid was a desirable starting point if external stimuli over the growth factor expression was going to be investigated.

The AqpZ-EGFP plasmid has KpnI and ApaI restriction enzyme recognition sites, flanking AqpZ gene. This was utilised to excise the AqpZ gene out of the vector and replaced it with the BMP2 sequence.

pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 is the last plasmid investigated, that was modified to include the BMP2 sequence. The original plasmid was designed by Shin and Noireaux and some of their design work included the truncation and modification at the N- and C-terminal of EGFP (Del6-229) ¹⁴. This is the minimal domain of EGFP that is obligatory for fluorescence¹⁵. Shin and Noireaux were also able to replace pTacI to the OR2-OR1-Pr promoter that is specific to sigma factor 70 in *E.coli*, add an untranslated-region (UTR1) containing a strong ribosome binding site for increased translation initiation efficiency, and include the addition of a transcriptional terminator, called T500, for the *E.coli* RNA polymerase. To extend the repertoire of cloning methodology in this project, Gibson technique was implemented to insert BMP2 sequence upstream/downstream of deGFP, creating two variants. The version of deGFP at the C-terminus of BMP2 was unsuccessful at further cell-free expression, and therefore, is not discussed here. On the other hand, deGFP at the N-terminus of BMP2 was pursuit further.

2.2 MATERIALS AND METHODS

2.2.1 GATEWAY CLONING

Gateway technology relies on site-specific-attachment (*att*) sites which contain a core 15 bp sequence on which the recombination occurs, and a surrounding region providing a binding site for the clonase enzymes¹⁶. In this recombinational cloning, the DNA fragment of interest is flanked by *attB1* and *attB2* sites (PCR product or an *attB* expression clone), whereas the Donor Vector has a toxic *ccdB* gene flanked by *attP1* and *attP2* sites. Integrase (Int) and the Integration Host Factor (IHF) proteins mediate the recombination of the *att* sites in what is called a BP reaction (Figure 2.1 A), resulting in an Entry Clone with the gene of interest now flanked by *attL1* and *attL2* sites. The following LR reaction (Figure 2.1 B), mediated by Int, IHF and Excisionase (Xis), involves the *attL1* and *attL2* sites on the newly generated Entry Clone from the BP reaction, and *attR1* and *attR2* sites on the Destination Vector, resulting in an Expression Clone. Here, gateway cloning was conducted on pCellFree plasmids to produce an N- and C-terminal BMP2 Expression Clone and an N- and C-terminal VEGF Expression Clone.

pCellFree_G03 (Addgene #67137), pCellFree_G04 (Addgene #67138), pCellFree_G05 (Addgene #67139) and pCellFree_G06 (Addgene #67140) were subject to GenElute Plasmid Miniprep (Sigma Aldrich), according to the manufacturer's instructions, prior to their use in Gateway cloning.

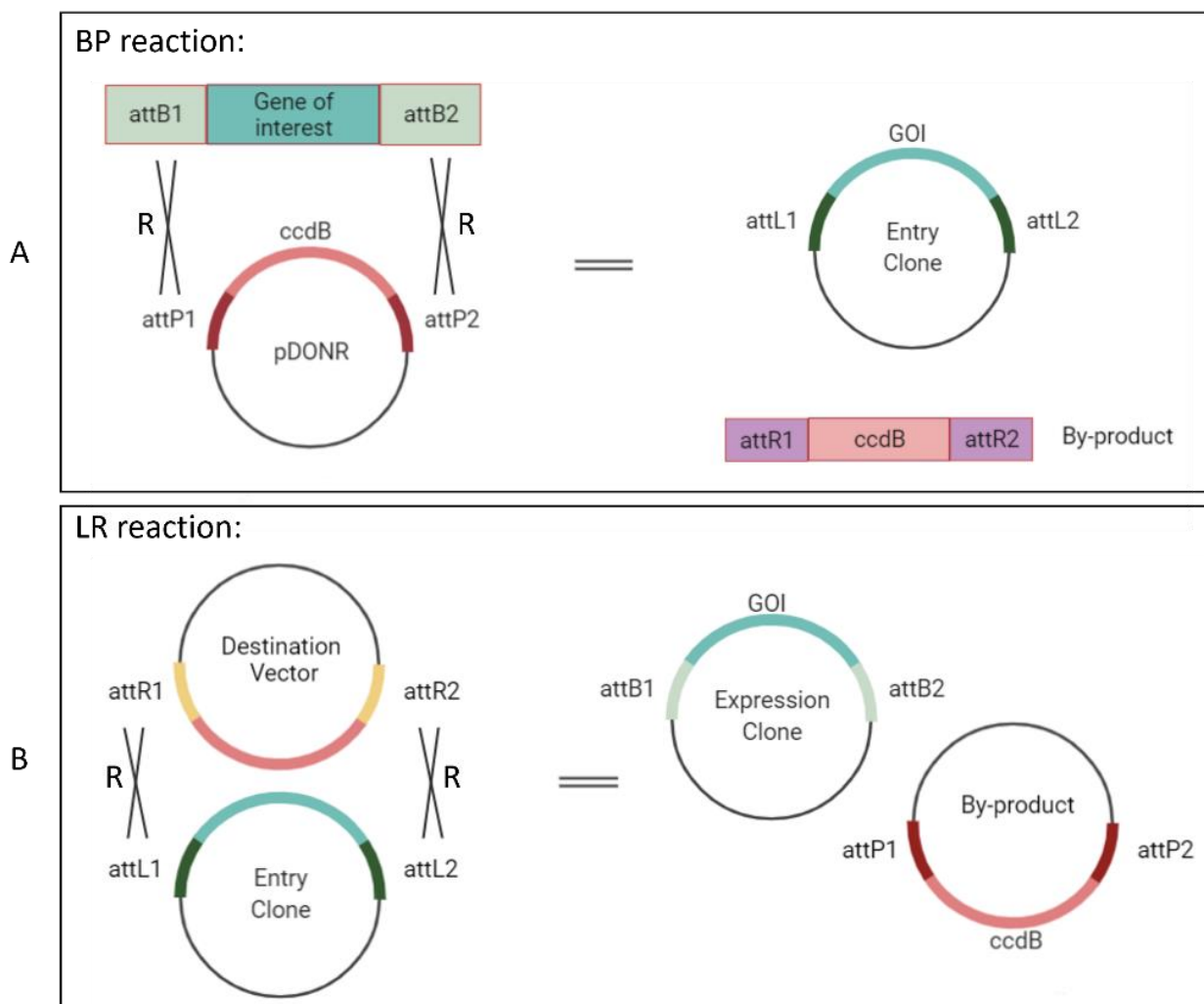


Figure 2.1 Schematic diagram showing gateway cloning using A: BP reactions and B: LR reactions. R= recombination event between corresponding att sites.

2.2.1.1 CREATING VEGF AND BMP2 SEQUENCES

VEGF and BMP2 lyophilised plasmids ordered from Eurofins, UK (see Appendix A) were dissolved in TE buffer (1 M Tris, 0.5 M EDTA, pH 8). One Shot TOP10 electrocompetent cells were transformed with 1 μ L DNA in an electroporation cuvette and electroporated at 25 μ F, 200 Ω , 2.5 kV. Non-salt yeast extract and nutrient broth (YENB) media was added to the transformed bacteria and incubated at 37°C, 225 rpm for 30 minutes before plating onto carbenicillin-containing agar plates. Glycerol stocks were created by mixing 50% sterile glycerol with bacterial culture in LB media, generating 25% glycerol bacterial

mixtures, then flash-freezing in liquid nitrogen. Polymerase Chain Reaction (PCR) was used to linearize the sequences at the following concentration conditions: 10 ng/ μ L VEGF pEX-A128 or BMP2 pEX-A258 DNA template, 10 μ M forward and reverse primers (see Appendix A), and Phusion Hot Start Flex 2X Master Mix (New England Biolabs). PCR settings were set according to Table 2.1. PCR products were loaded onto 1% DNA agarose gel (Section 2.2.2.2) and separated at 100 V, 100 mA for 1 hour. The bands corresponding to the VEGF fragment and BMP2 fragment were cut out from the gel and purified by QIAquick gel extraction kit (Qiagen) following the manufacturer's instructions. The expected size of full length linearised VEGF and BMP2 are 627 bp and 1242 bp respectively.

<i>Steps</i>	<i>Times and Temperature</i>	<i>Cycles</i>
Initial denaturation	1 min at 98°C	x1
Denaturation	15 sec at 98°C	x34
Annealing	30 sec at 55°C +/-3°C	
Extension	30 sec at 72°C	
Final extension	5 min at 72°C	x1
Hold	∞ at 4°C	x1

Table 2.1 Thermocycler programme for generating linear VEGF and BMP2 sequences for Gateway cloning.

2.2.1.2 BP REACTION

Linearized BMP2 and VEGF sequences were added separately as 50 fmol to 150 ng/ μ L pDONR in TE buffer, pH 8. Gateway BP Clonase II enzyme mix was added as 2 μ L to the samples and vortexed twice, before overnight incubation at 25°C. The reaction was stopped by the addition of Proteinase K and incubation at 37°C for 10 minutes. One Shot TOP10 *E.coli* cells were transformed with 2 μ L of the BP Clonase sample mixture in the

same manner as previously described in Section 2.2.1.1, and plated out on zeocin-containing agar plates.

2.2.1.3 LR REACTION

The BMP2 containing Entry Clones created from BP reactions were added as 150 ng to 150 ng of pCellFree_G03 and pCellFree_G04, whereas VEGF containing Entry Clones were mixed with pCellFree_G05 and pCellFree_G06 Destination Vectors. This reaction was carried out in the same manner as BP reaction, but with 2 μ L of the Gateway LR Clonase II enzyme mix, and the One Shot TOP10 *E.coli* cells were plated onto carbenicillin-containing plates after transformation.

2.2.2 ENGINEERING A TWO-PLASMID SYSTEM

A two-plasmid system was generated to induce the expression of BMP2-EGFP fusion. For this, a plasmid containing σ 28 sequence on a pBEST-Luc vector was used to allow the transcription of another plasmid (AqpZ-EGFP), which was modified to replace AqpZ sequence with the BMP2 sequence (see Appendix A). Class II restriction endonucleases are the most common restriction enzymes used for molecular cloning due to their high DNA specificity and ATP-independent cleaving¹⁷. The mechanism of action consists of several steps, with the first being the exclusion of water as the enzyme interacts non-specifically with the phosphate backbone of DNA. This is followed by sequence scanning until the recognition site is found, then additional hydrogen bonds are formed creating a tight enzyme:DNA complex. In the presence of Mg^{2+} , the phosphodiester bond on each strand becomes cleaved, generating either a blunt end or a staggered cut, also known as sticky ends. DNA fragments from different sources containing complementary ends can be joined together *via* a new phosphodiester bond through the use of a DNA ligase.

2.2.2.1 RESTRICTION DIGESTION

For the restriction digestion of AqpZ-EGFP plasmid, two enzymes were used: KpnI HF (New England Biolabs) and ApaI (New England Biolabs) to remove AqpZ sequence and

linearise the plasmid. The same restriction enzymes were used to create the sticky ends on the corresponding insert BMP2 sequence. 2 μL of 424 ng/ μL ApqZ-EGFP plasmid or 348 ng/ μL BMP2 sequence was mixed with 5 μL of 10x CutSmart Buffer, 42 μL nuclease free water, and 1 μL ApaI restriction enzyme, then incubated at 25°C for 1 hour. 1 μL of KpnI HF was added and incubated at 37°C for 1 hour. The resulting DNA was subject to gel electrophoresis. The incubation was also changed to 15 minutes or an overnight digestion to compare effectiveness.

2.2.2.2 GEL ELECTROPHORESIS

Small DNA fragments were applied to a 2% gel, whereas higher molecular weight fragments were applied to a 1% gel. A 2% agarose gel was made by dissolving 1.2 g of agarose in 60 mL of Tris-acetate-EDTA (TAE) buffer (40 mM Tris, 20 mM acetic acid, 1mM EDTA) and stained with 6 μL SYBR Safe dye (Thermo Fisher). DNA samples were mixed with Purple Loading Dye (New England Biolabs) as a 6:1 volume ratio. For smaller wells, 5 μL of either 1 kbp PLUS DNA ladder (New England Biolabs) or 100 bp DNA ladder (New England Biolabs) was loaded into the first wells, followed by 15 μL reaction volumes. 100 V, 100 mA was applied to the gel for 1 hour. The DNA fragments corresponding to the correct weights after restriction digestion were cut out from the agarose gel and subjected to the gel extraction protocol (Qiagen).

2.2.2.3 LIGATION AND TRANSFORMATION

Purified DNA vector and insert were ligated in several different ratios according to the formula:

$$\frac{(ng\ of\ vector) \times (kb\ size\ of\ insert)}{(kb\ size\ of\ vector) \times (insert:vector\ ratio)} = ng\ of\ insert\ required$$

Equation 1. Calculation of the amount of insert required for ligation ratio.

The ratios of insert to vector which were applied for this ligation included 1:1, 2:1, 3:1, 4:1 as well as a negative control of no insert. The reactions were set to 10 μL volumes

which included 5 μL of vector, the ratio corresponding to the volume of insert, 1 μL T4 ligase, 1 μL T4 buffer and the remainder of the volume was filled with nuclease free water. The ligation reactions were gently mixed by pipetting and incubated overnight at 4°C. The resulting reaction of each sample was added as a volume of 2 μL to Top10 *E.coli* cells and an electroporation protocol was followed. Alternatively, the 2 μL was added to JM109 *E.coli* cells and a heat shock protocol followed after. For the electroporation protocol, the cells/sample mixture was transferred into a pre-chilled electroporation cuvette (Sigma Aldrich) and electroporated at 100 Ω , 25 μF . SOC outgrowth media (New England Biolabs) was added immediately after and incubated at 37°C for 30 minutes before plating onto agar plates. In the heat shock protocol, the cells/sample mixture was returned on ice for 30 minutes, after which the cells were heated to 42°C for 20 seconds. The mixture was returned to ice for additional 2 minutes and then 450 μL of SOC outgrowth media (New England Biolabs) was added and incubated at 37°C for 30 minutes before plating onto agar plates.

2.2.3 GIBSON ASSEMBLY

Gibson Assembly is a cloning procedure that allows the joining of DNA fragments *via* homologous overlapping ends, without the need for restriction enzymes¹⁸. The 5'-3' exonuclease activity present in the Gibson Assembly Master Mix creates a single stranded 3' overhang on each DNA fragment. Since the overhangs are complementary on the separate DNA fragments, annealing can occur. The now double-stranded, annealed DNA is extended by DNA polymerase and sealed by DNA ligase.

Here, Gibson Assembly was utilized to create an N-terminal deGFP-BMP2. First, BMP2 fragment was designed to include complementary overlapping nucleotides. The same overlapping sequence was then designed onto the existing pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid (see Appendix A).

2.2.3.1 GENERATION OF LINEAR DNA FRAGMENTS

For N-terminal deGFP-BMP2, two DNA fragments had to be created. PCR was used to generate the fragments, where the reactions contained the following: 10 μ M of the fragment specific forward and reverse primers (see Appendix A), Phusion Hot Start II High-Fidelity PCR Master Mix and 10 ng/ μ L DNA template (pEX-A258-BMP2 for Fragment 1 BMP2 and pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 for Fragment 2). PCR reactions were set up as shown in Table 2.2, on Geneflow Sensoquest basic thermocycler. The generated PCR products were run on the 1% DNA agarose gel and the right sized fragments were cut out and purified using the QIAquick gel extraction kit (Qiaagen) following the manufacturer's instructions. The expected BMP2 fragment with overlap was calculated at 402 bp and the linear vector segment at 2584 bp.

<i>Steps</i>	<i>Times and Temperature</i>	<i>Cycles</i>
Initial denaturation	1 min at 98°C	x1
Denaturation	15 sec at 98°C	x34
Annealing	30 sec at 65°C +/-6°C	
Extension	30 sec at 72°C	
Final extension	5 min at 72°C	x1
Hold	∞ at 4°C	x1

Table 2.2. Thermocycler programme for generating linear DNA fragments for Gibson assembly.

2.2.3.2 GIBSON ASSEMBLY TECHNIQUE

The Assembly Mixture at 15 μ L (320 μ L 5x ISO buffer, 0.64 μ L of 10 U/ μ L T5 exonuclease, 20 μ L of 2 U/ μ L Phusion polymerase, 160 μ L of 40 U/ μ L Taq DNA ligase, and water to 1.2 mL) was thawed on ice, and 5 μ L of the DNA to be assembled was mixed in. The DNA fragments were in equimolar amounts. The Gibson Assembly aliquot and DNA fragments mixture was incubated at 50°C for 1 hour, and then transformed into

One Shot TOP10 *E.coli* cells as described previously. Transformed bacteria was plated onto carbenicillin agar plates.

2.2.3.3 COLONY POLYMERASE CHAIN REACTION

To confirm the success of Gibson Assembly, the colonies formed on carbenicillin agar plate were subject to Colony PCR. Single colonies were re-suspended in 20 μ L sterile MiliQ water and then heated to 95°C for 5 minutes. They were then centrifuged at 13 500 rpm for 1 minute and 5 μ L of the supernatant was used as the DNA template for PCR. Reference plates were also created for easy retrieval of the successful clones. The rest of the PCR mix comprised of OneTaq 2x Master Mix with Standard Buffer (New England Biolabs), 10 μ M of forward (forward primer of Frag 1 BMP2) and reverse (reverse primer of Frag 2) primers, and nuclease-free water. PCR settings were set according to Table 2.3, and run on Geneflow Sensoquest basic thermocycler. PCR products were run on 1% DNA agarose gel and the successful clones were recognised by their corresponding molecular weight of the DNA band shown. The representative clones were then taken from the reference plate and grown as a standard starter culture, which was then subject to GenElute Plasmid Miniprep (Sigma Aldrich).


<i>Steps</i>	<i>Times and Temperature</i>	<i>Cycles</i>
Initial denaturation	2 min at 94°C	x1
Denaturation	30 sec at 94°C	 x30
Annealing	30 sec at 75°C	
Extension	30 sec at 72°C	
Final extension	5 min at 72°C	x1
Hold	∞ at 4°C	x1

Table 2.3 Thermocycler programme for colony PCR.

All of the generated plasmids in this chapter have been subjected to nucleotide sequencing by Eurofins Genomics, for further sequence confirmation.

All of the relevant plasmids obtained or created in this thesis are outlined in Table 2.4.

Plasmid name	Promoter	Relevant gene sequence (N'-C')	Cloning technique/ Origin
pBEST-OR2-OR1-Pr-UTR1-deGFP-T500	P70	deGFP-Stop	Noireaux group
pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500	P70	deGFP-TEV-BMP2-6xHis-Stop	Gibson Assembly
pCellFree_G03	T7	8xHis-EGFP-HRV3C-attR1-ccdB-CamR-attR2-Stop	Alexandrov group
pCellFree_G03+BMP2	T7	8xHis-EGFP-HRV3C-attB1-BMP2-attB2-Stop	Gateway cloning
pCellFree_G04	T7	attR1-ccdB-CamR-attR2-HRV3C-EGFP-8xHis-Stop	Alexandrov group
pCellFree_G04+BMP2	T7	attR1-BMP2-attR2-HRV3C-EGFP-8xHis-Stop	Gateway cloning
pCellFree_G05	T7	8xHis-mCherry-HRV3C-attR1-ccdB-CamR-attR2-Stop	Alexandrov group
pCellFree_G05+VEGF	T7	8xHis-mCherry-HRV3C-attB1-VEGF-attB2-Stop	Gateway cloning
pCellFree_G06	T7	attR1-ccdB-CamR-attR2-HRV3C-mCherry-8xHis-Stop	Alexandrov group
pCellFree_G06+VEGF	T7	attB1-VEGF-attB2-HRV3C-mCherry-8xHis-Stop	Gateway cloning
Pr1-T7RNAP	P70	T7RNAP-Stop	Noireaux group
pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500	P70	σ 28-Stop	Noireaux group
pTar-AqpZ-EGFP	P28	AqpZ-EGFP-Stop	Gifted by Dr Angelique Coutable
pTar-BMP2-EGFP	P28	BMP2-TEV-EGFP-Stop	Restriction enzymes

Table 2.4 Summary of the plasmids obtained and generated in this thesis. The table information includes the full plasmid name, promoter of the plasmid, the relevant sequence of the plasmid elements and the method of cloning or the origin of the given plasmid.

2.3 RESULTS AND DISCUSSION

2.3.1 PCELLFREE

To create input DNA encoding for the growth factors BMP2 and VEGF, and for it to be functional with CFES, pCellFree DNA plasmids were chosen. These plasmids have been designed by Gagoski *et al.*, and are described as destination vectors with Species Independent Translation Initiation Sequence (SITS), allowing the mediation of expression in both prokaryotic and eukaryotic *in vitro* TX-TL systems¹². These destination vectors have an N- or a C-terminal EGFP/mCherry/sfGFP as reporter protein sequences for easy expression tracking, and an 8xHis-tag, for further purification if desired. The destination vectors are Gateway cloning compatible with the first step being the linearisation of VEGF and BMP2 sequences along with attB sites ready to create Entry Clones (Figure 2.2).

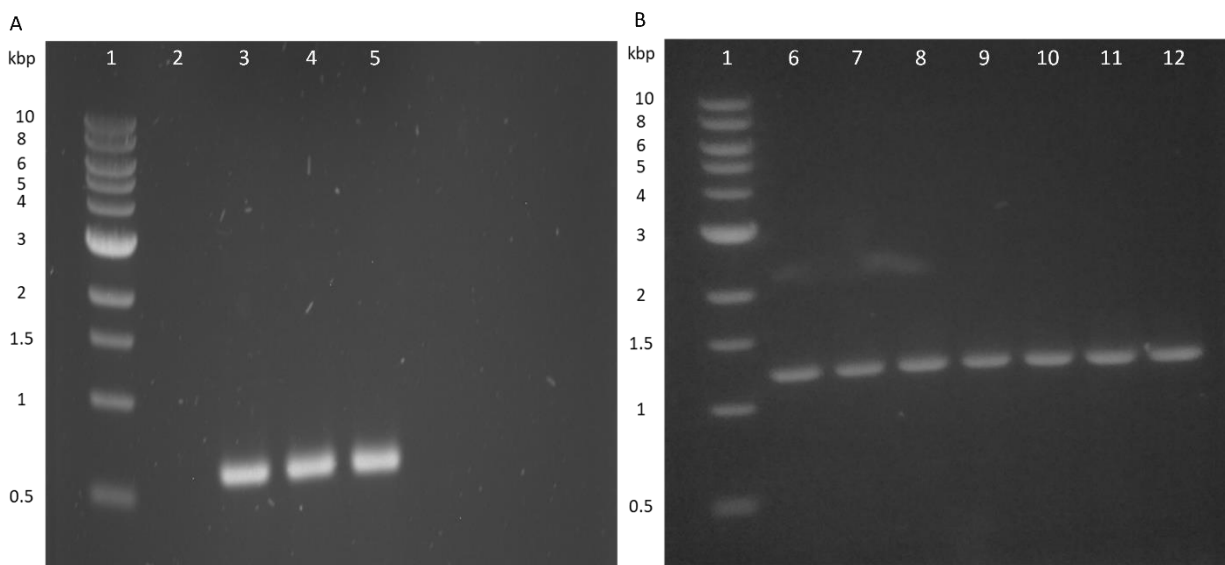


Figure 2.2 Agarose gel electrophoresis of PCR products. A: VEGF linearisation. B: BMP2 linearisation. Lane 1 indicates 1 kb DNA ladder with successfully separated size markers. Lane 2 is a PCR product from a reaction with no template DNA. Lane 3 contains PCR products when the annealing temperature was set to 56°C. Lane 4 and 5 are PCR products of the annealing temperature 56.7°C. Lanes 6-12 contain BMP2 linearisation reactions with the annealing temperature set to 55°C. The gels were subject to UV-light ($\lambda = 302$ nm) to take the image.

Linear fragments of VEGF with attB sites are depicted by the thick bands present in Figure 2.2 A when the annealing temperature for PCR was set to 56°C and 56.7°C, and

Chapter 2 - Plasmid Design

the bands correspond to ~627 bp. The successful BMP2 fragments correspond to the bands present at around 1242 bp in Figure 2.2 B. The DNA fragments were carefully excised and purified before being inserted into pDONR vectors in the BP reaction. The generated Entry Clones were transformed into competent cells and assessed by colony count. pDONR vectors contain zeocin resistance gene past attP sites, and a chloramphenicol resistance gene along with the suicide *ccdB* gene flanked by attP sites. No colonies were noted in the chloramphenicol counterselection agar plate, but multiple colonies were present on the zeocin agar plate, from which they were selected for overnight culture and later DNA mini-prep. The purified Entry Clones and pCellFree plasmids were subject to LR reaction and the success was assessed by colony count on selection plates as well as DNA gel electrophoresis (Figure 2.3).

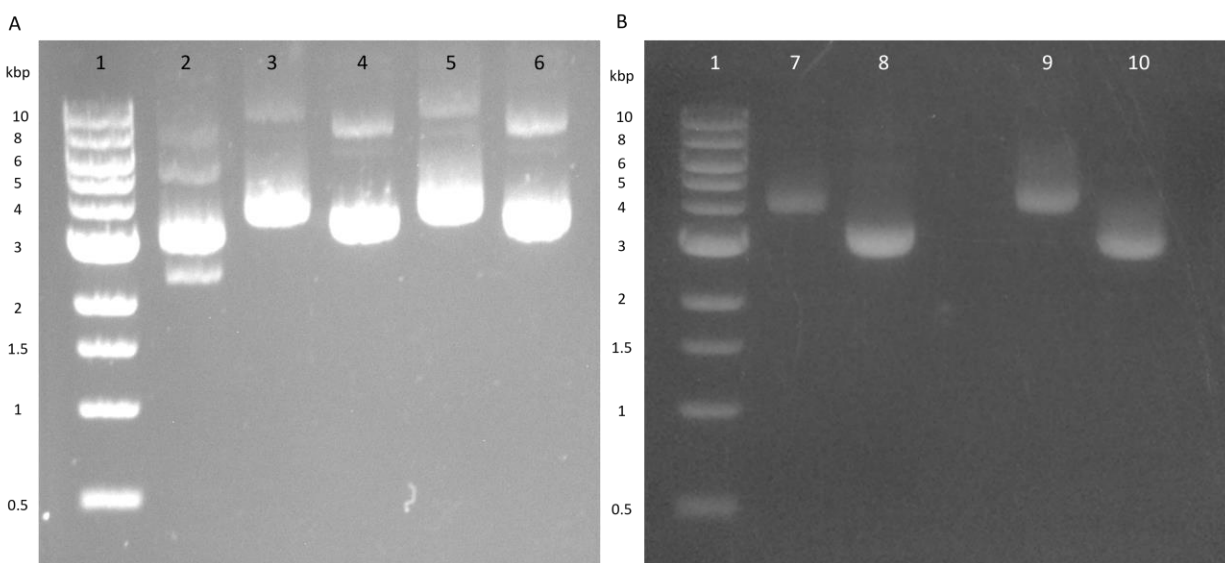


Figure 2.3 Agarose gel electrophoresis of pCellFree and Gateway cloned pCellFree plasmids. A: EGFP containing plasmids. B: mCherry containing plasmids. Lane 1 = 1 kbp DNA ladder. Lane 2 = BMP2 pDONR Entry Clone. Lane 3 = pCellFree_G03, Lane 4 = pCellFree_G03+BMP2, Lane 5 = pCellFree_G04, Lane 6 = pCellFree_G04+BMP2, Lane 7 = pCellFree_G05, Lane 8 = pCellFree_G05+VEGF, Lane 9 = pCellFree_G06, Lane 10 = pCellFree_G06+VEGF. The gels were subject to UV-light ($\lambda = 302$ nm) to take the image.

pCellFree plasmids contain chloramphenicol resistance gene and a suicide *ccdB* gene between the attR sites. Once the attL sites of the insert recombined with attR sites of the

pCellFree vector, the resulting plasmid only had resistance against ampicillin that is present outside of the att sites. The LR reaction mix transformed into competent cells gave rise to colonies present on the carbenicillin plates and no colonies on the agar plate containing chloramphenicol. Figure 2.3 shows the DNA originating from these colonies, against unmodified pCellFree plasmids. The nucleotides including the *ccdB* lethal gene and the chloramphenicol resistance gene that are inbetween the attR sites in pCellFree plasmids amount to 1603 bp. This larger segment was replaced with either a much smaller 627 bp VEGF or a 1242 bp BMP2 sequence, indicating that the Gateway cloning was successful.

2.3.2 TWO-PLASMID SYSTEM

A two-plasmid system was created to add modularity and an engineered control over the expression of the growth factors. The pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 plasmid was built by Shin *et al.* to expand the toolkit of regulatory calibre of the endogenous *E.coli* transcription machinery¹⁴. The plasmid with pTar promoter that would be transcribed in response to the presence of *E.coli* sigma factor 28 was obtained in a form of an Aquaporin Z (AqpZ)-EGFP from Dr Angelique Coutable (School of Biochemistry, University of Bristol)¹³. This expression plasmid was modified to replace the AqpZ gene with a BMP2 sequence, by using restriction enzyme digestion. The result of the insert switch was analysed by DNA gel electrophoresis. Figure 2.4 A shows there is a weight difference between single cut 2450 bp pEX-A128 BMP2 vector, and double-digested plasmid with ApaI and KpnI HF enzymes, however, the 371 bp BMP2 fragment with overhang sticky ends is not visible in the agarose gel. The pEX-A128 BMP2 plasmid was mixed again with the two restriction enzymes and incubated at either 15 minutes or overnight (Figure 2.4 B). In the agarose gel in Figure 2.4 B, faint bands just below 0.4 kbp (encircled) represent the correctly digested BMP2 sequence with overhang sticky ends.

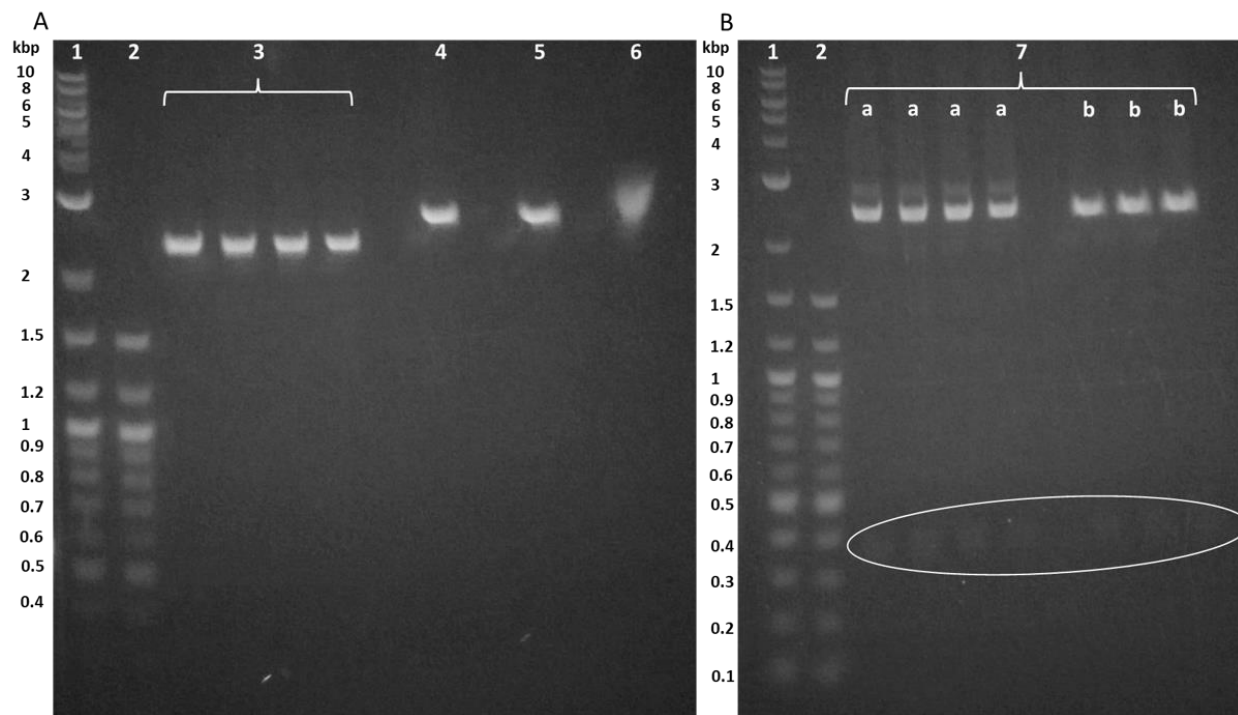


Figure 2.4 Agarose gel electrophoresis after restriction digestion of BMP2 with ApaI and KpnI HF restriction enzymes. Lane 1 = 1 kb PLUS DNA ladder. Lane 2 = 100 bp DNA ladder. Lane 3 = ApaI and KpnI HF cut BMP2. Lane 4 = ApaI only cut BMP2. Lane 5 = KpnI HF only cut BMP2. Lane 6 = uncut BMP2. Lane 7 = ApaI and KpnI HF cut BMP2, Lane 7a = 15 minutes digestion reaction of BMP2, Lane 7b = an overnight digestion reaction of BMP2. The gels were subject to UV-light ($\lambda = 302 \text{ nm}$) to take the image.

The small bands corresponding to cut BMP2 were excised and extracted for ligation with the vector at various ratios. The vector was also subjected to double restriction digestion with ApaI and KpnI HF, so that complementary sticky ends would be generated for the ligation with BMP2. Once TOP10 competent cells were transformed with the ligation reactions, and colonies were grown on agar plates, the DNA from single colonies was analysed using agarose gel electrophoresis for successful ligation and transformation as shown in Figure 2.5. The restriction digestion efficiency of AqpZ-EGFP is much easier to assess because AqpZ gene fragment is 693 bp, almost double the base pairs of BMP2.

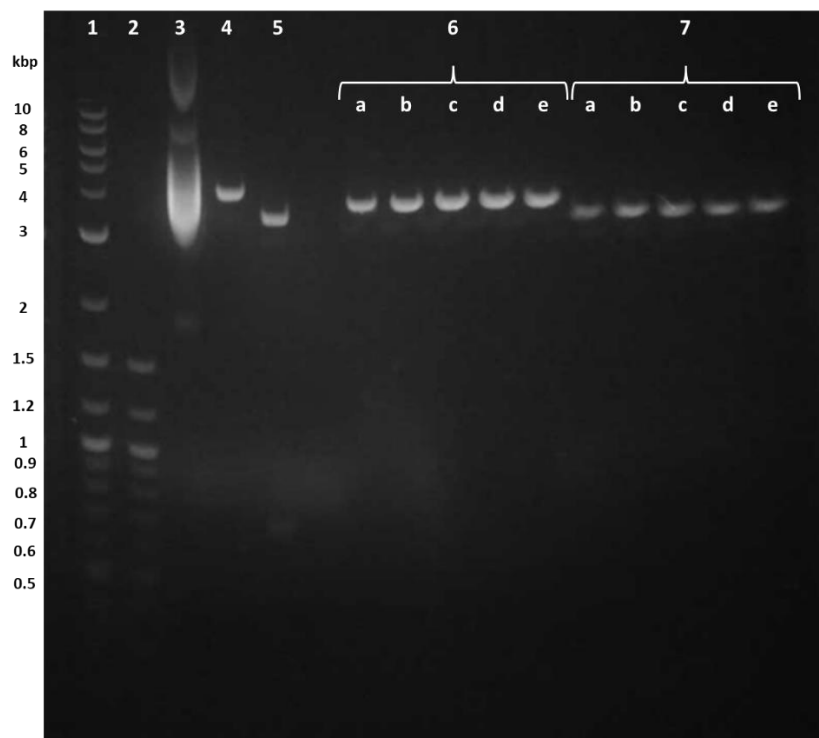


Figure 2.5 Agarose gel electrophoresis after vector digestion and ligation with BMP2 insert. Lane 1 = 1 kb PLUS DNA ladder. Lane 2 = 100 bp DNA ladder. Lane 3 = Uncut, supercoiled AqpZ-EGFP vector. Lane 4 = ApaI cut vector. Lane 5 = ApaI and KpnI HF cut vector. Lane 6 = ApaI digestion of DNA extracted from ligation colonies, Lane 6a = colony 1 from ligation reaction performed at 1:1 ratio, Lane 6b = colony 2 from ligation reaction performed at 1:1 ratio, Lane 6c = colony 3 from ligation reaction performed at 2:1 ratio, Lane 6d = colony 4 from ligation reaction performed at 2:1 ratio, Lane 6e = colony 5 from ligation reaction performed at 3:1 ratio. Lane 7 = ApaI and KpnI HF digestion of DNA extracted from ligation colonies at the same ratios as Lane 6a-e. The gel was subject to UV-light ($\lambda = 302$ nm) to take the image.

There is an obvious weight change between ApaI cut vector (Figure 2.5, Lane 4) and ApaI/KpnI HF double cut vector (Figure 2.5, Lane 5), which most likely correspond to the 693 bp AqpZ gene removal. The ligation reactions between the restriction-digested EGFP vector and BMP2 sequence have been conducted at various ratios, and all of the ratios tested by gel electrophoresis have been successful as shown in Figure 2.5, Lane 6 and 7. When the ligated DNA from those colonies was subjected to ApaI restriction enzyme (Figure 2.5, Lane 6), a 3906 bp linearised plasmid was generated, as seen on the gel. This was 371 bp heavier than the DNA subjected to both ApaI and KpnI HF enzymes (Figure 2.5, Lane 7) because of the successfully ligated-in BMP2 sequence. These bands

also appear lighter than linearised AqpZ-EGFP, because the BMP2 insert is shorter than AqpZ sequence. It is also important to note that both the double-digested vector before the ligation and the double-digested expression plasmid from the ligation colonies are the same weight as seen in Lanes 5 and 7 of Figure 2.5.

2.3.3 DEGFP-BMP2

The pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid has been used as a reference and calibrator input DNA for endogenous *E.coli* CFES, and therefore, was modified to contain BMP2 sequence in hopes of high expression yields in CFES. Firstly, both deGFP vector and BMP2 sequence were linearised and flanked by overlapping sequences for Gibson cloning. The generation of those fragments was achieved by PCR and assessed by gel electrophoresis (Figure 2.6).

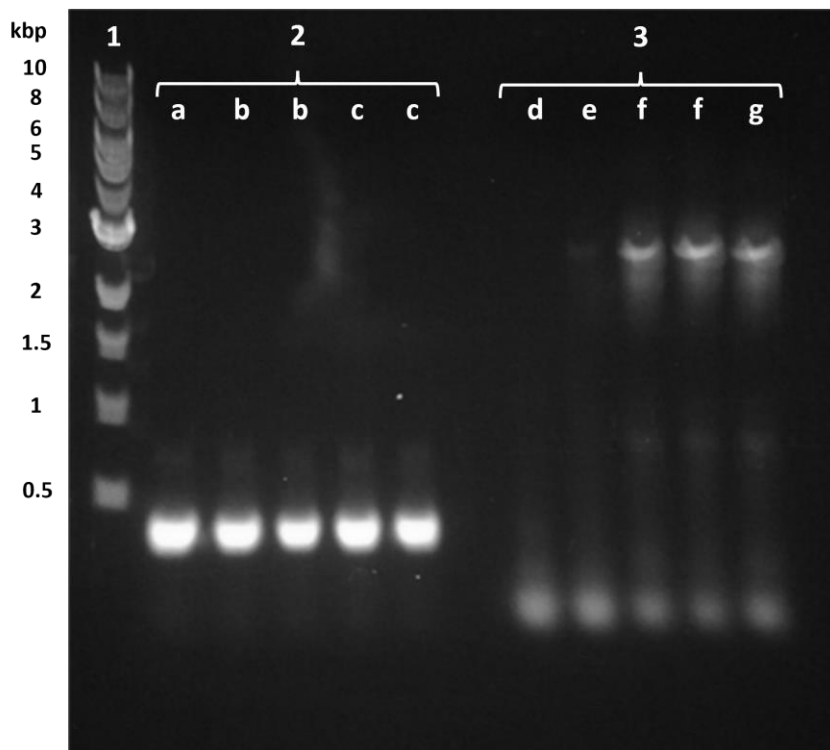


Figure 2.6 Agarose gel electrophoresis of PCR reactions for the amplification of Gibson fragments. Lane 1 = 1 kb DNA ladder. Lane 2 = BMP2 fragment amplified by PCR with the annealing step set to temperatures at a = 59.6°C, b = 60.5°C, c = 61.7°C. Lane 3 = deGFP vector fragment amplified by PCR with the annealing step set to temperatures at d = 65.7°C, e = 67.1°C, f = 68.3°C, g = 69.5°C. The gel was subject to UV-light ($\lambda = 302$ nm) to take the image.

Chapter 2 - Plasmid Design

With the expected size of BMP2 with overlap region being 402 bp, the PCR reactions have been successful at the tested range of the annealing temperatures in Lanes 2. These five bands have been cut out of the agarose gel and purified for further experiments. The expected size of the deGFP vector with overlap region, totals to 2584 bp, which can be seen in the PCR reactions with the annealing step set to 68.3°C and 69.5°C. The top bands shown in Lane 3f and 3g in Figure 2.6 were excised and purified. The temperature below 68.3°C was likely to be too low for the primers to anneal to the vector fragment and give rise to amplification. After Gibson Assembly of the purified BMP2 and deGFP vector fragments with overlaps, the resulting mixture was used in transformation of TOP10 cells. The colonies that grew on carbenicillin agar plates were screened using PCR with the forward and reverse BMP2 fragment primers. This combination of primers would test for the presence of the inserted BMP2 sequence within the vector (Figure 2.7).

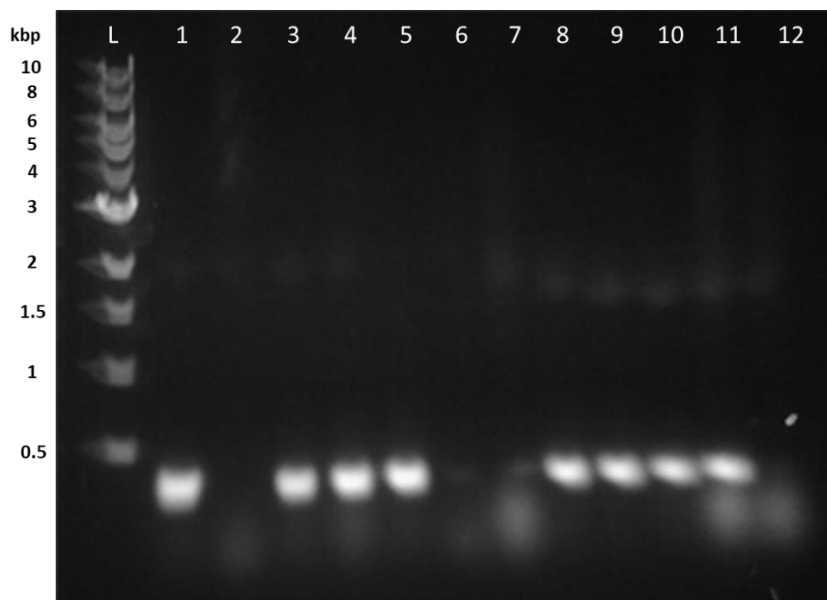


Figure 2.7 Agarose gel electrophoresis following colony PCR of Gibson assembled N-terminal deGFP-BMP2. L = 1 kb DNA ladder. Lane 1-12 = corresponding colony number used in the amplification of a fragment in N-terminal deGFP-BMP2 plasmid. The gel was subject to UV-light ($\lambda = 302$ nm) to take the image.

Figure 2.7 reveals the amplification of the 402 bp BMP2 fragment within the DNA originating from the colonies numbered 1, 3, 4, 5, 8, 9, 10 and 11. No strong bands were

Chapter 2 - Plasmid Design

visible from PCR amplification using DNA from colonies 2, 7, 8 and 12, suggesting the BMP2 insert might not be present in the deGFP vector. The colonies with correctly ligated inserts were grown overnight and subjected to DNA mini-prep, where the purified DNA was used for CFES experiments.

2.4 CONCLUSIONS AND FURTHER WORK

In this chapter, several expression plasmids were successfully generated. Beginning with the Gateway cloning of a series of pCellFree plasmids to generate chimeras consisting of mCherry with VEGF and EGFP with BMP2. The growth factor sequences were added to the N- or the C-terminals of the reporters within the plasmids. The use of restriction enzymes was utilised with a positive outcome to produce a plasmid with BMP2-EGFP sequence, under a pTar promoter which is recognisable by sigma factor 28. The ligations at numerous ratios proved high yielding in bacterial colonies with correctly assembled plasmids. The last type of plasmid designed and created here was the deGFP-BMP2 fusion. The BMP2 sequence was inserted downstream of deGFP on pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 plasmid by applying the Gateway cloning methodology. The success of molecular cloning adapted for the three types of plasmids was measured by numerous ways, including a positive selection system with a lethal gene, restriction mapping, as well as colony screening with PCR. The array of constructed plasmids facilitates a vast number of applications for the transcription and translation in an *E.coli* endogenous system. Some of them include low-complexity monitoring of expression through the presence of the reporter protein, isolation and purification possibility of the synthesised protein due to the attached histidine-tag, TEV protease/HRV 3C protease site would allow for cleavage of the fusion-tag from the growth factor, if desired, and a two-plasmid system could provide a stepping-stone into the development of external stimuli for CFES. A part of future work could include designing molecular breadboards with genetic toggle switches¹⁹, oscillators²⁰, or logic gates²¹ to better mimic the *in-vivo* release kinetics of growth factors. The successfully constructed plasmids described in this chapter were subjected to rounds of cell-free expression to optimise fusion protein yields.

2.5 BIBLIOGRAPHY

1. Kahn, T. W., Beachy, R. N. & Falk, M. M. Cell-free expression of a GFP fusion protein allows quantitation in vitro and in vivo. *Current Biology* 7, (1997).
2. Kapust, R. B. & Waugh, D. S. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8, 1668–1674 (1999).
3. Smith, D. B. & Johnson, K. S. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* 67, 31–40 (1988).
4. Bornhorst, J. A. & Falke, J. J. Purification of proteins using polyhistidine affinity tags. *Methods in Enzymology* 326, 245–254 (2000).
5. Raran-Kurussi, S. & Waugh, D. S. The Ability to Enhance the Solubility of Its Fusion Partners Is an Intrinsic Property of Maltose-Binding Protein but Their Folding Is Either Spontaneous or Chaperone-Mediated. *PLoS ONE* 7, (2012).
6. Raran-Kurussi, S., Cherry, S., Zhang, D. & Waugh, D. S. Removal of Affinity Tags with TEV Protease. *Methods in Molecular Biology* 1586, 221–230 (2017).
7. Müller-Hill, B. *The lac Operon: a short history of a genetic paradigm*. (Walter de Gruyter, 1996).
8. Hansen, L. H., Knudsen, S. & Sørensen, S. J. The Effect of the lacY Gene on the Induction of IPTG Inducible Promoters, Studied in Escherichia coli and Pseudomonas fluorescens. *Current Microbiology* 36, 341–347 (1998).
9. Tabor, S. & Richardson, C. C. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci U S A* 82, 1074–1078 (1985).

10. de Boer, H. A., Comstock, L. J. & Vassert, M. The tac promoter: A functional hybrid derived from the trp and lac promoters. *Proc. NatL Acad. Sci. USA* 80, 21–25 (1983).
11. Shin, J. & Noireaux, V. An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synthetic Biology* 1, 29–41 (2012).
12. Gagoski, D. *et al.* Gateway-compatible vectors for high-throughput protein expression in pro- and eukaryotic cell-free systems. *Journal of Biotechnology* 195, 1–7 (2015).
13. Coutable, A. *et al.* Preparation of tethered-lipid bilayers on gold surfaces for the incorporation of integral membrane proteins synthesized by cell-free expression. *Langmuir* 30, 3132–3141 (2014).
14. Shin, J. & Noireaux, V. Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of Biological Engineering* 4, (2010).
15. Li, X. *et al.* Deletions of the Aequorea victoria Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *Journal of Biological Chemistry* 272, 28545–28549 (1997).
16. Hartley, J. L., Temple, G. F. & Brasch, M. A. DNA Cloning Using In Vitro Site-Specific Recombination. *Genome Research* 10, 1788 (2000).
17. Williams, R. Restriction Endonucleases: Classification, Properties, and Applications. *Molecular Biotechnology* 23, 225–243 (2003).
18. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* 6, 343–347 (2009).
19. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in Escherichia coli. *Nature letters* 403, 339–342 (2000).

Chapter 2 - Plasmid Design

20. Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature letters* 456, 516–520 (2008).
21. Lehr, F.-X. *et al.* Cell-free prototyping of AND-logic gates based on heterogeneous RNA activators. *ACS Synthetic Biology* 8, 2163–2173 (2019).

CHAPTER 3. OPTIMISATION OF THE CELL-FREE EXPRESSION SYSTEM

3.1 INTRODUCTION

Cell-free protein synthesis is an emerging method for fast and high-throughput generation of proteins from genetic information. Lysates from many species are currently utilised as cell-free extract sources, including those from rabbit reticulocytes¹, wheat-germ embryo², *Leish tarentolae*³ or *E.coli*⁴. It is also possible to achieve cell-free protein synthesis by using systems made of expressed, purified and reconstituted transcription-translation (TX-TL) machinery components, such as PURExpress, which tend to be high-yielding⁵. Here, the cell lysate was produced in-house from *E.coli* following an established protocol published by Sun *et al.*⁶, and outputs from plasmids generated in Chapter 2 were compared to a commercially available S30 *E.coli* extract⁷ as well as the reconstituted *E.coli* system, PURExpress⁵. The crude extract originally developed by Sun *et al.*, is based on BL21-Rosetta 2 *E.coli* strain with the OmpT and Lon protease genes deleted, but rare tRNAs genes added, maximising protein production⁸. Other efficiency- and yield-increasing modifications to the crude extract selected by Sun *et al.* included the utilisation of 3-phosphoglyceric acid (3-PGA) as the energy source over creatine phosphate or phosphoenolpyruvate, the use of Mg- and K-glutamate over Mg- and K- acetate, and the removal of 2- mercaptoethanol, a potent reducing agent⁶. Bead-beating as a method of cell lysis was chosen over sonication or homogenisation to reduce running costs with no negative impact on protein production⁹. Plasmid pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (abbreviated to pdeGFP) has been specifically designed to be recognised and transcribed by the extract published by Sun *et al.*¹⁰. Although numerous plasmids have been developed by the Noireaux group for the endogenous *E.coli* TX-TL system, pdeGFP has been the highest yielding fluorescent reporter so far¹¹. This feature contributed to the reason why pdeGFP was used as a benchmark for extract calibrations as well as CFES yield comparisons with newly designed plasmids as discussed in this chapter.

One of the main aims of this project was to design and produce a functional plasmid, encoding for a growth factor with a reporter protein, that is compatible with a cell-free

expression system. Two growth factors were of particular interest; vascular endothelial growth factor (VEGF) which is a major signalling molecule in angiogenesis,¹² and bone morphogenetic protein 2 (BMP2), an inducer of osteogenesis¹³. Both are vital during bone repair, when the granulation tissue full of differentiating mesenchymal stem cells is formed and new vasculature is created¹⁴. Although both growth factors were considered at the beginning, in the interest of time, plasmids developed later in the project only focused on BMP2 sequence cloning. As soon as the plasmid containing a growth factor and a fluorophore sequence was designed and produced (Chapter 2), it was tested for compatibility with CFES (Chapter 3). However, when the plasmid was not generating expression yields, another set of plasmids was undergoing development to ensure a successful candidate for expression in CFES was in place. Some of the input DNAs were only fully optimised into protein-producing plasmids with CFES, after a working system has already been developed for other plasmids.

The first plasmids modified with either VEGF or BMP2 sequence at the N- or C-terminal of the fluorophore were pCellFree (Chapter 2). Despite them being recognised by numerous species due to species independent translation initiation sequence (SITS), they were optimised by Gagoski *et al.* using mostly *Leish tarentolae* and *E.coli* based CFES¹⁵. Another method of producing a growth-factor-fluorophore fusion protein *via* CFES, encompassed a two-plasmid system. This method utilised pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 as the primary circular DNA recognised by the endogenous *E.coli* RNA polymerase¹⁶. Once sigma factor 28 was expressed, the transcription of the BMP2-EGFP fusion on pTar-BMP2-EGFP plasmid could be initiated. This additional level of expression control was advantageous, especially since bone healing progression is largely influenced by strict timing of the growth factor release, as well as its concentration and location¹⁷. The final plasmid that was measured for protein production yields in CFES was the pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500 (pdeGFP-BMP2). A major

benefit of this plasmid was that it contained the already established high yielding reporter compatible with the endogenous *E.coli* TX-TL.

Hydrogels, a general term encompassing 3-dimensional polymer networks, are able to retain large amounts of water without dissolving. This property provides a soft tissue-like environment for cells, where they can exist in a preferred 3D space¹⁸. In addition, their plasticity allows for filling of irregularly shaped locations, their porosity and permeability for load uptake and nutrient diffusion, and the hydrogel's properties can usually be tuneable¹⁹. These qualities render hydrogels useful for tissue engineering, and in this project, the reason why hydrogels were investigated is because they could act as a potential carrier of growth-factor-producing CFES to safely deliver BMP2 or VEGF to more precise locations in bone fractures. Although CFES are mostly solution based, revolutionary new chassis for CFES been recently investigated and so far include cell-mimics²⁰, paper²¹ and various hydrogels^{22, 23} which are discussed in more detail in Chapter 1. Here, agar, agarose and in-house designed bioink were brought to focus for the purpose of housing growth-factor-producing CFES. Both agar and agarose have been one of the most highly cell-free-expressing hydrogels tested by Whitfield *et al.*²³. A bioink designed by the Perriman group consisting of Pluronic F127 and alginate, was shown to be cytocompatible, with a high resolution extrudability²⁴. The Perriman group was able to utilise the F127-alginate bioink for tissue engineered cartilage and bone prints. To push the frontiers of this bioink's engineering, it was compelling to extend it to cell-free protein synthesis. In this chapter, CFES with pdeGFP-BMP2 as input DNA were mixed with the three different hydrogels and the fluorescence output from the gels were measured. Preliminary diffusion of the fluorescent growth factor from the gels was also evaluated by comparing the levels of fluorescence between the gel and the surrounding liquid in the well.

3.2 MATERIALS AND METHODS

3.2.1 IN-HOUSE CELL-FREE EXPRESSION SYSTEM

Crude cell extracts were prepared from BL21-Rosetta2 *E.coli*, grown in 5 L Erlenmeyer flasks containing 660 mL 2xYT+P media with 34 mg/mL chloramphenicol at 37°C and 220 rpm until OD600 reached 1.2. The cells were collected by centrifugation of 1 L batches at 5000 g, 4°C, for 15 minutes. The bacterial pellet was subject to washing and resuspension in buffer S30A (14 mM Mg-glutamate, 60 mM K-glutamate, 50 mM Tris, 2mM DTT, pH 7.7) twice, with centrifugation at 5000 g, 4°C for 12 minutes. After the final centrifugation step, the pellet was immersed in liquid nitrogen for -80°C storage. To each gram of wet pellet mass, 0.9 mL of buffer S30A was added, vortexed and returned on ice. Then, a total of 5 g of 0.1 mm diameter beads was intermittently added to each gram of wet pellet mass and vortexed, forming a thick paste. The bead-cell solution was transferred into bead-beating tubes and bead beating was carried out at 46 rpm for 1 minute total per tube. Micro chromatography filter apparatuses were assembled and the extract was separated from pellet and beads by centrifuging at 6000 g, 4°C for 5 minutes. Only correctly bead-beaten extracts were processed further (Figure 3.1). The separated supernatant was then subject to further centrifugation at 12 000 g, 4°C for 10 minutes. Digestion of any remaining nucleic acids was done by the released endogenous exonucleases when the supernatant was incubated at 37°C, 180 rpm, for 80 minutes. Final centrifugation at 12 000 g, 4°C for 10 minutes was performed before subjecting the cell extract to dialysis using 10k MWCO dialysis cassettes in buffer S30B (14 mM Mg-glutamate, 60 mM K-glutamate, ~5 mM Tris, 1 mM DTT, pH 8.2) at 4°C for either overnight (Extracts 1, 3, 4) or 3 hours (Extract 2). Total protein concentration was determined using a Bradford assay standard, after which the extracts were aliquoted at 30 mg/mL and flash-frozen in liquid nitrogen for long term storage at -80°C.



Figure 3.1. Visual quality assessment of viable extract. Left hand side: poorly processed extract appears as turbid. Right hand side: the extract has distinct layers and a clear supernatant; therefore, it has been correctly bead-beaten.

3.2.1.1 AMINO ACID SOLUTION

The solution was made using RTS Amino Acid Sampler kit which provides all 20 amino acids at 168 mM, except for leucine at 140 mM. Once thawed, the amino acids were added together in the following order Ala, Arg, Asn, Asp, Gln, Glu, Gly, His, Ile, Lys, Met, Phe, Pro, Ser, Thr, Val, Trp, Tyr, Leu, Cys to make a final composition of 6 mM, with leucine at 5 mM. The solution was aliquoted at 26 μ L, 52 μ L and 500 μ L, flash-frozen and stored at -80°C .

3.2.1.2 ENERGY SOLUTION

The energy solution was made to have a final composition of 700 mM HEPES (pH 8), 21 mM ATP, 21 mM GTP, 12.6 mM CTP, 12.6 mM UTP, 2.8 mg/mL tRNA, 3.64 mM CoA, 4.62 mM NAD, 10.5 mM cAMP, 0.95 mM folinic acid, 14 mM spermidine and 420 mM 3-PGA. The final solution was aliquoted into 7 μ L and 150 μ L volumes, flash-frozen and stored at -80°C .

3.2.1.3 TRANSCRIPTION-TRANSLATION (TX-TL) REACTION SETUP

TX-TL reactions were made up from the crude cell extract, energy solution, amino acid solution and DNA. The ratio was 75% buffer and 25% DNA. In a 10 μ L reaction volume, the final conditions were as follows: 8.9-9.9 mg/mL crude cell extract, 1.5 mM amino acids, apart from leucine at 1.25 mM, 30 mM 3-PGA, 1 mM spermidine, 0.75 mM cAMP, 0.33 mM NAD, 0.068 mM folinic acid, 0.26 mM CoA, 0.2 mg/mL tRNAs, 50 mM HEPES, 1.5 mM ATP, 1.5 mM GTP, 0.9 mM CTP, 0.9 mM UTP, 2% PEG-8000, and a range of Mg-glutamate and K-glutamate concentrations, depending on the extract calibration results. The reactions were set up on ice and briefly vortexed after each addition, then transferred to a 384-well plate for fluorescence readings using Biotek Synergy Neo2 plate reader (BioTek). The temperature was set to 29°C or 37°C, and the stage shaking was set to ‘orbital’ for 10 seconds before taking measurements at each timepoint. Time-lapse was normally set to 16 hours with measurements collected every 15 minutes. The excitation was at 479 ± 20 nm and emission at 520 ± 20 nm for pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (Addgene #40019) plasmid DNA, excitation at 485 ± 10 nm and emission at 525 ± 20 nm for EGFP containing plasmids (pCellFreeG03 and pCellFreeG04) and excitation at 579 ± 10 nm with emission at 616 ± 20 nm for mCherry containing plasmids (pCellFreeG05 and pCellFreeG06).

3.2.2 PUREXPRESS REACTIONS SETUP

PURExpress is a commercially available transcription-translation system initially developed by Takuya Ueda, that is reconstituted from *E.coli* TX-TL purified components²⁵. Since the system is made out of recombinant purified elements, the amount of present RNases, exonucleases or proteases is kept to a minimum, giving rise to high yields of expression. A T7 RNA polymerase is present in the PURExpress kit, therefore, in this project, this system was used for the transcription and expression of pCellFree plasmids, which contain a T7 promoter.

All components of the PURExpress *In Vitro* Protein Synthesis kit (New England Biolabs) were thawed on ice. Solution A (5 μL) was mixed with 3.75 μL Solution B, along with 0.33 μL 40U/ μL RNase inhibitor (New England Biolabs) and template DNA with nuclease-free water, making the volume up to 13.1 μL . The mixture was transferred to a 384-well plate for fluorescence readings using Biotek Synergy Neo2 plate reader. The temperature was set to 37°C, and the stage shaking was set to ‘orbital’ for 10 seconds before taking measurements at each timepoint.

3.2.3 *E. COLI* S30 EXTRACT REACTIONS SETUP

The *E. coli* S30 Extract System is another commercially available TX-TL coupled kit. This system, however, is based on an *E. coli* extract, which was first purified by Geoffrey Zubay, and later modified to increase stability of expressed proteins⁷. This kit allows the transcription and translation of DNA originating from a plasmid or lambda vectors. Here, the system was used as a comparison method for the expression of both the plasmids under the T7 promoter and plasmids under the bacteriophage lambda promoter.

The components of the *E. coli* S30 Extract System (Promega) were thawed on ice. The 10 μL reactions constituted 4 μL S30 Premix Plus solution, 3.6 μL T7 S30 extract solution, along with nuclease-free water and template DNA. The mixture was transferred to a 384-well plate for fluorescence readings using Biotek Synergy Neo2 plate reader. The measurements were taken at 37°C with 10 seconds stage shake before each timepoint.

3.2.4 HYDROGEL CHASSIS

An alginate-Pluronic F127 bioink, agar and agarose hydrogels were prepared at 2 and 4% w/v for 1:1 v/v and 1:3 v/v dilutions with CFES. For agar and agarose, 20 mg and 40 mg of each powder (G-Biosciences) were mixed with 1 mL MiliQ water and vortexed before placing at 95°C for 15 minutes to dissolve. After dissolving, the temperature was lowered and maintained at 60°C. For 2% w/v bioink, 120 mg sodium alginate (Sigma Aldrich) and 260 mg Pluronic F127 (Sigma Aldrich) was mixed with 1 mL ice-cold diethylpyrocarbonate (DEPC) treated water, whereas for a 4% w/v, 240 mg sodium

alginate and 520 mg Pluronic F127 was mixed with 1 mL ice-cold DEPC water. Thorough hand-mixing was performed with a pipette tip, and the 5 mL eppendorfs containing the bioinks were transferred to 4°C. CFES reactions with either pdeGFP or pdeGFP-BMP2 as input DNA, were prepared as stated in section 2.1, but with corrected calculations for 20 μL as final volume. Using a positive-displacement pipette, 10 μL or 15 μL CFES reactions were added to 10 μL or 5 μL gels to achieve the 1:1 v/v and 1:3 v/v dilutions. The mixing between agar/agarose and CFES was carried out at 60°C and the mixing between the bioink and CFES was performed on ice, then transferred into a 384-well plate or into top right section of a 96-well plate. For diffusion experiments, the gels were set before adding 30 μL DEPC water to each well. Bioink containing reaction had 100 mM CaCl_2 supplemented DEPC water for cross-linking. The plate reader was set to ‘area scanning’, where the well of a 96-well plate was divided into 9 separate sections, each being scanned every 30 minutes at 29°C for the diffusion experiments.

3.3 RESULTS AND DISCUSSION

3.3.1 DEGFP

pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 is a plasmid optimised by Noireaux group, to contain specialised sequences for enhanced protein production in *E.coli* based TX-TL¹⁰. The modifications were specifically designed to regulate the endogenous *E.coli* TX-TL mechanism, therefore, in this project, the plasmid was chosen as a model input DNA for the in-house BL21-Rosetta2 CFES. Following the protocol by Sun *et al.* for crude cell extract production, the concentration of extract 1, 2, 3 and 4 was 32.7 mg/mL, 33.0 mg/mL, 29.9 mg/mL and 29.0 mg/mL respectively, as determined by the Bradford assay standard⁶. Each extract must be calibrated with an optimal concentration of ions, which are essential for a multitude of enzymatic activities and interactions in TX-TL, including the stability of ribosomal units, tRNAs and rRNAs²⁶. The most prevalent anions in TX-TL being acetate or glutamate, and the cations being magnesium (Mg^{2+}) and potassium (K^+), with both cations working in synergy^{27, 28}. The CFES reactions were set up as mentioned in Section 3.2.1.3 with Mg-glutamate in the range of 0-6 mM and K-glutamate in the range of 20-140 mM, and the fluorescence output was measured using the plate reader (Figure 3.2). Over time, four extracts were generated to keep up with the stock demands for CFES reactions. The ion calibration was judged by the amount of deGFP produced as a function of maximum fluorescence intensity measured over time. The ion calibration of the extract 1 in Figure 3.2 showed the most optimal concentration of Mg-glutamate to be 2 mM and 80 mM K-glutamate. For extract 2, it was 3 mM Mg-glutamate and 20 mM K-glutamate. The highest output in extract 3 was generated from 5 mM Mg-glutamate and 40 mM K-glutamate, whereas in extract 4, it was found that 4 mM Mg-glutamate and 80 mM K-glutamate resulted in highest fluorescence. Although all four extract outputs cannot be directly compared due to different gain settings, extract 4 is highest yielding, as will be demonstrated further. This could be attributed to improved

experience and stricter selection of higher-quality aliquots at the last protocol steps, as shown in Figure. 3.1.

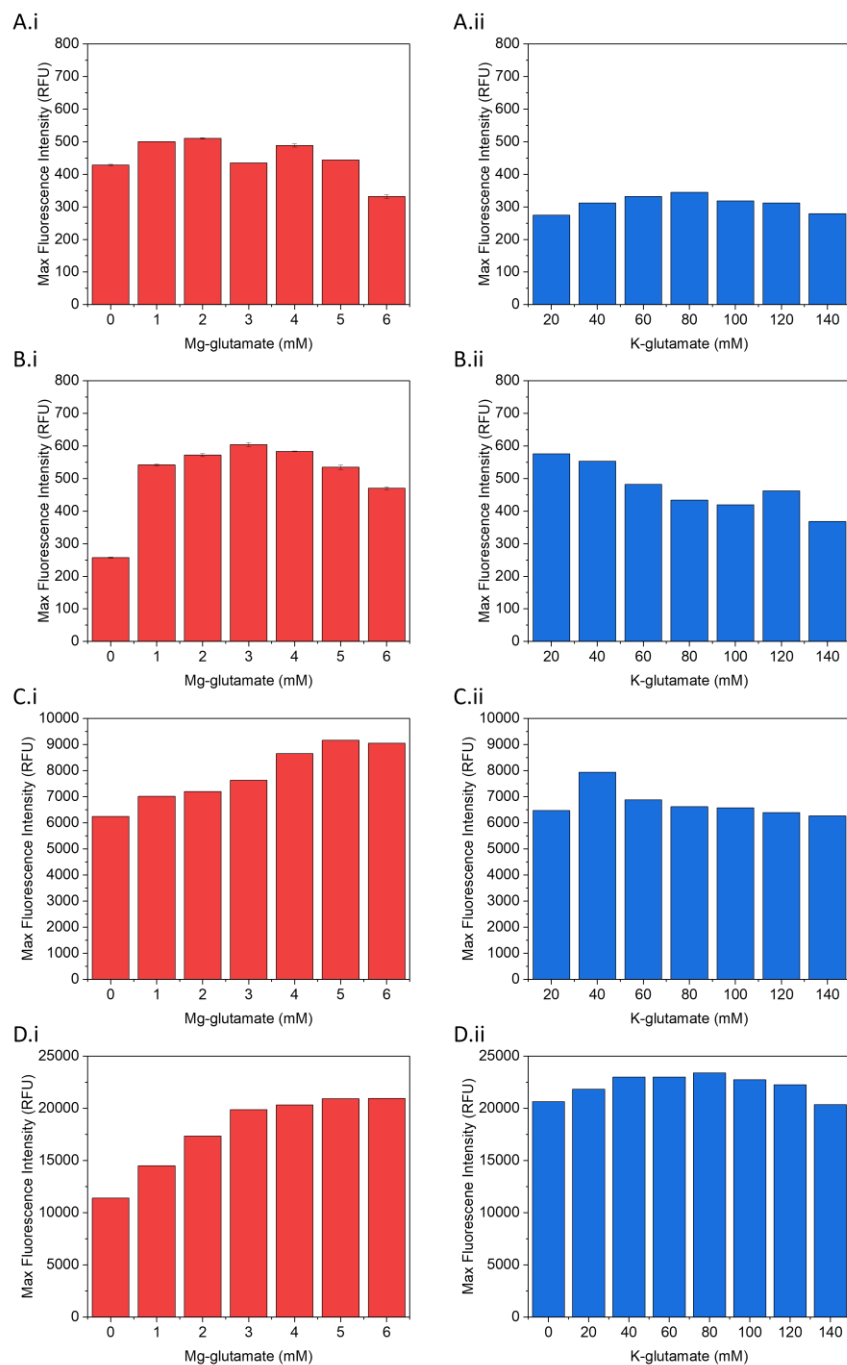


Figure 3.2 Calibration of extracts 1-4 (A-D) with Mg-glutamate (i) and K-glutamate (ii). Maximum fluorescence intensity after a 16 hours reaction run at 29°C. BioTek Synergy Neo2 set at gain 50 for A and B, and at gain 70 for C and D. 1nM pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 as input DNA. Excitation= 485 ± 10 nm, emission= 520 ± 20 nm.

3.3.2 pCELLFREE

After successful Gateway cloning of pCellFree vectors with BMP2 and VEGF sequences (Chapter 2, Section 2.3.1), the new pCellFree+BMP2/VEGF plasmids were used as input DNA with the in-house CFES (Figure 3.3).

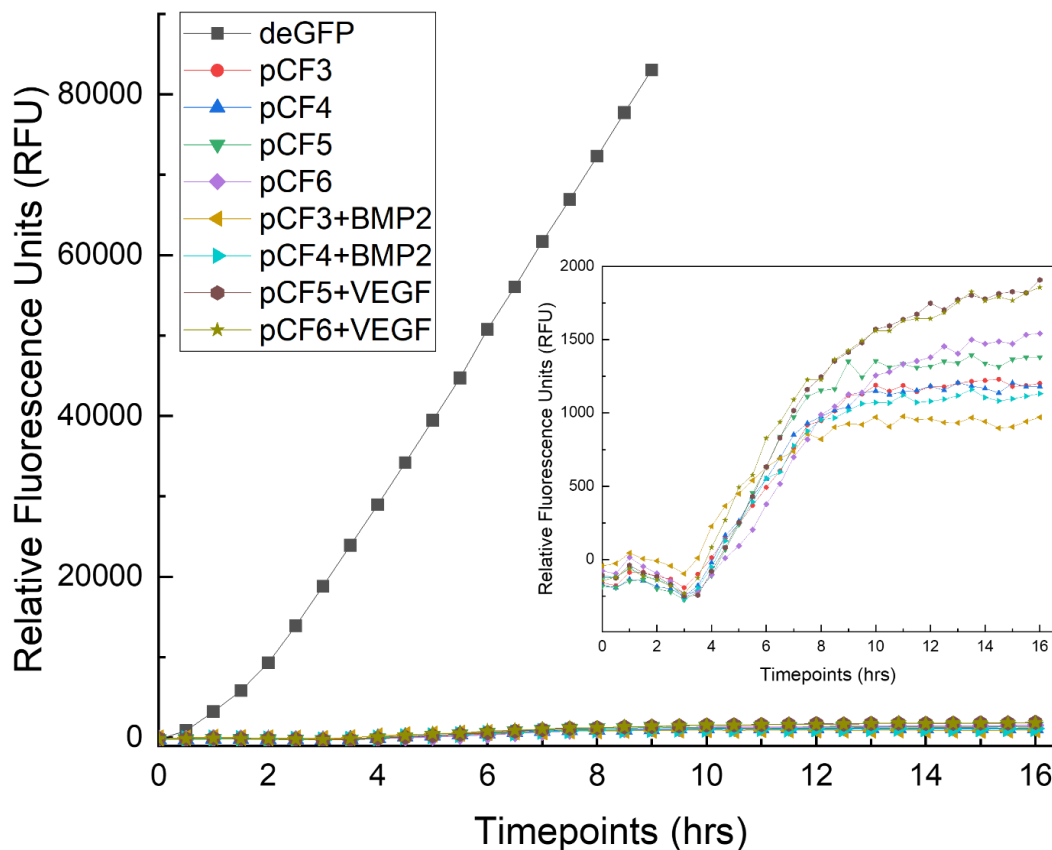


Figure 3.3 Kinetics of deGFP (black), EGFP (red and blue), mCherry (green and purple), EGFP-BMP2 fusion (yellow and cyan) and mCherry-VEGF fusion (brown and olive) expression from 9 nM input plasmids, pdeGFP, pCF and pCF+VEGF/BMP2. Background (CFES with DEPC water) subtracted. Insert = deGFP excluded. Excitation= 485 ± 10 nm, emission= 520 ± 20 nm. Temperature= 29°C . Gain= 90. Extract 1.

Figure 3.3 shows a clear, steady increase in fluorescence over time when pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 (pdeGFP) was used as input DNA (black), then at the 9-hour time point, 84207 RFU was reached generating an ‘overflow’ reading due to a saturated detector. The way around this is a sample dilution or lowering the gain setting in order to continue taking the measurements. However, this is applicable to future

experiments, as the time-lapse experiment could not be paused. The fluorescence readings from reactions containing unmodified and modified pCellFree input DNA are significantly lower than those from deGFP. An increase in fluorescence over time from pCellFree reactions can be seen on a different y -axis scale as reproduced in the insert of Figure 3.3. All eight pCellFree plasmids give rise to a similar kinetics profile, where a steady increase is noted just before the 4th hour after reaction initiation, before some of the profiles reaching a plateau around the 8th hour. The fluorescence intensity that is reached after 16 hours in the pCellFree reactions is between ~1000-2000 RFU, and interestingly, due to an experimental error, mCherry fluorescence follows the same kinetic profile as EGFP, yet the diffraction gratings were set for EGFP excitation and emission. Although the fluorescence from extract only reactions have been already subtracted, this similarity between mCherry and EGFP suggests that the fluorescence output from pCellFree seen in Figure 3.3 is background noise. The reason as to why pCellFree input DNA did not generate fluorescent proteins in the in-house CFES, can be explained in the plasmid design. All of the pCellFree vectors contain a T7 promoter in front of the open reading frame, and considering the in-house CFES is based on the endogenous *E.coli* RNAP, no transcription of the pCellFree DNA can occur until the extract is supplemented with T7 RNAP.

The next experiment focused on just one plasmid, pCellFreeG03+BMP2, to bring about TX-TL using the in-house CFES supplemented with T7 RNAP (Figure 3.4). The varying amount of T7 RNAP added to the extract at four different plasmid concentrations was tested. Throughout the 55 hours monitoring, the plots in Figure 3.4 did not appear to resemble the sigmoidal curves that are normally observed in CFES reactions. No significant trend has unravelled across the range of the T7 RNAP tested here. The fluorescence values are very minimal and therefore, inconsequential, and no increase in values is noted when the input DNA concentration is raised from 1 nM, across 9 nM and 15 nM to 30 nM. The result of 30 nM pCellFreeG03+BMP2 with no addition of T7 RNAP

(Figure 3.4 D, black line), shows an increase in fluorescence within the first couple of hours before reaching plateau after 10 hours. It is intriguing that although the T7 RNAP is required for the transcription of the plasmid, the more typical sigmoidal curve is observed when T7 RNAP was not included in the reaction mixture along with 30 nM DNA. However, as mentioned before, the yield was underwhelming and should be considered as not above the background signal.

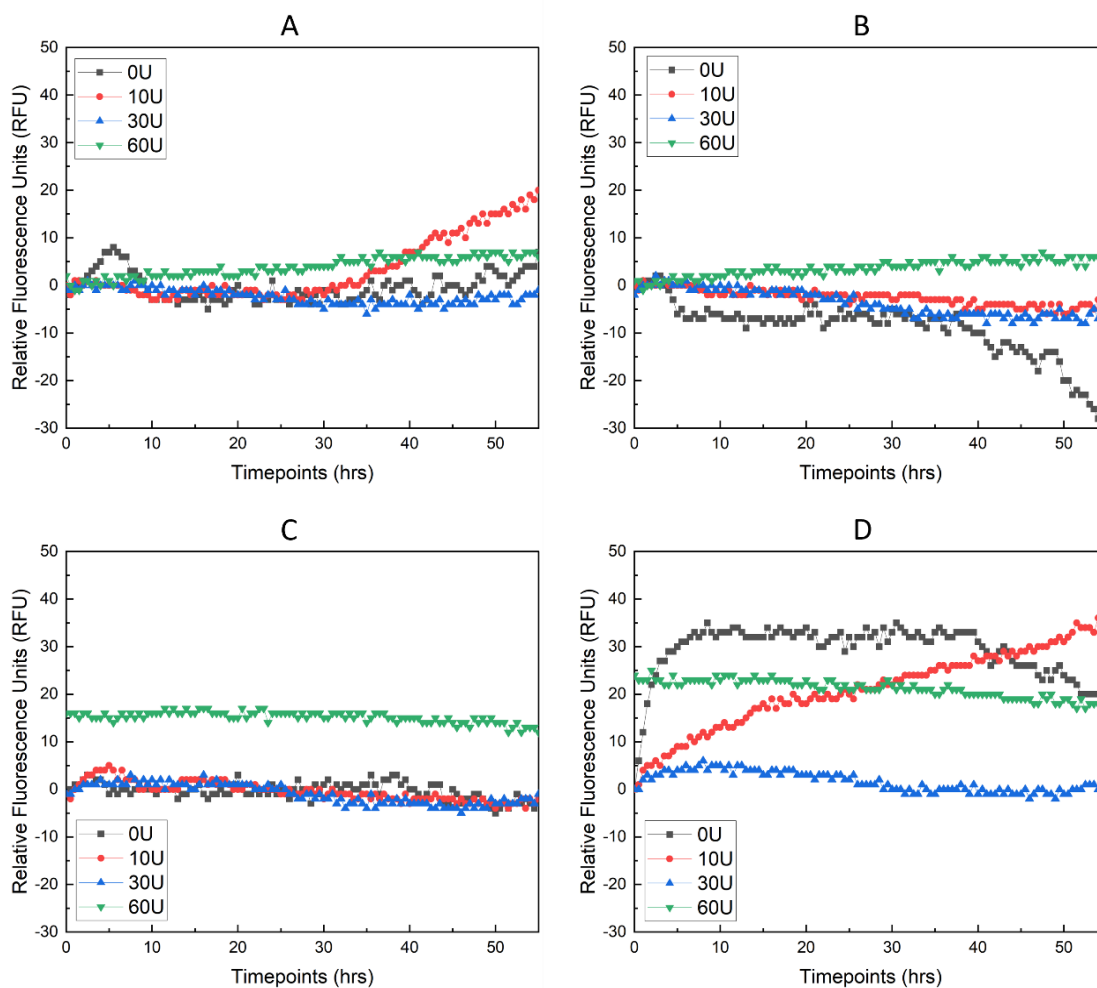


Figure 3.4 Kinetics of EGFP-BMP2 expression in the in-house CFES, originating from pCFG03+BMP2, supplemented by 0 U T7 RNAP (black), 10 U T7 RNAP (red), 30 U T7 RNAP (blue) and 60 U T7 RNAP (green). Four concentrations of input DNA: A: 1 nM, B: 9 nM, C: 15 nM, D: 30 nM. Background reactions with no DNA have been subtracted from the readings. Excitation = 485 ± 10 nm, emission = 520 ± 20 nm. Temperature = 29°C . Gain = 50. Extract 1.

Although the tested range of T7 RNAP did not bring about the production of EGFP-BMP2 fusion from pCellFreeG03+BMP2, it was noted that the plasmids also contained a lac operator. The lactose (lac) operon consists of a series of genes that are crucial in the transport and metabolism of lactose in *E.coli*²⁹. There are three genes, *lacZ*, *lacY* and *lacA*, under the influence of a singular promoter P_{lac} ³⁰. These genes code for β -galactosidase, galactoside permease and galactoside O-acetyltransferase respectively. An operator site, O_{lac} , is situated between P_{lac} and *lacZYA*, and it is at this site that the lac repressor attaches to and minimises transcription by RNA polymerase in the absence of an inducer³¹. The lac repressor is encoded by the *lacI* gene, under the regulation of a separate promoter P_{lacI} , both located upstream from the rest of the lac operon elements. The presence of isopropyl- β -D-thiogalactopyranoside (IPTG) causes a conformational change in the repressor, leading to its unbinding from the operator site and allowing for the full activity of the RNA polymerase³². Usually, the location of the lac operon, on the expression plasmid is affiliated with the promoter for the gene of interest. This offers the possibility of regulating the expression of the gene of interest, *via* the addition of IPTG. Unusually, the location of the lac promoter on pCellFree plasmids is downstream of the open reading frame. Despite this, the addition of IPTG to the CFES containing pCellFreeG03+BMP2 was tested. The results are summarised as a heatmap in Figure 3.4, where the reactions which yielded the highest levels of EGFP-BMP2 are denoted in blue. The lowest fluorescence outputs, where the fluorescent fusion production was not necessarily successful, appears as a red colour on the heatmap. The addition of T7 RNAP at the units explored in Figure 3.5 did not increase the protein production, just as shown in Figure 3.4. In fact, even when 5U of T7 RNAP was added, the maximum fluorescence output decreased by half, when compared to reactions with no T7 RNAP. Perhaps the addition of another polymerase enzyme could give rise to a competition for the binding to a promoter, with an already existing *E.coli* RNAP in the extract. Since the competition

between the enzymes may result in short ‘on’ and ‘off’ binding states, the transcription rate would be negatively impacted, leading to lower expression yield.

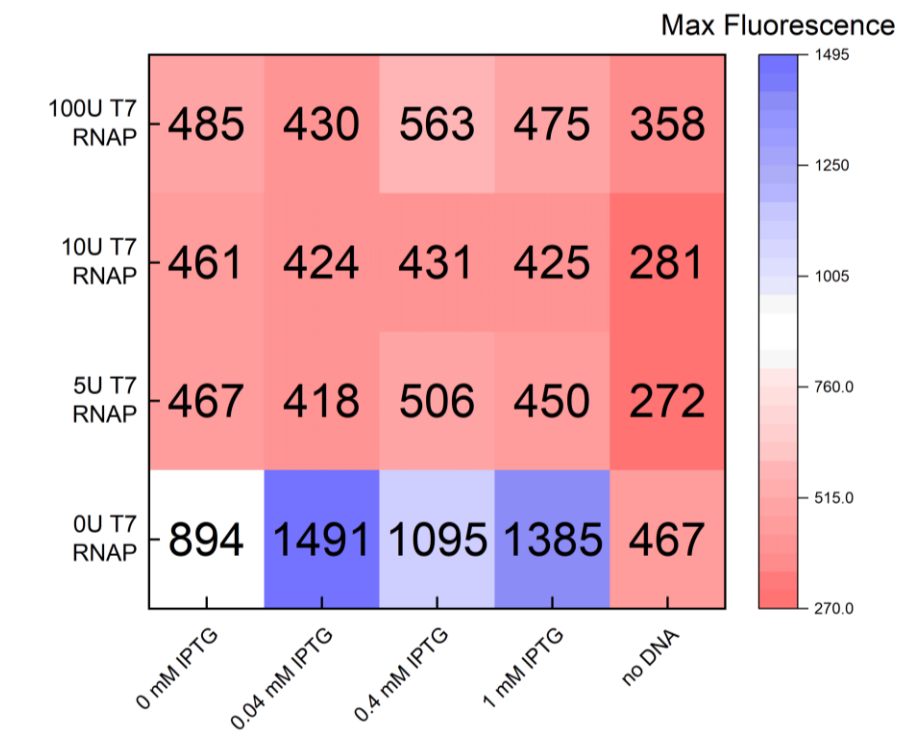


Figure 3.5 Heatmap of the addition of IPTG and T7 RNAP at a range of concentrations to CFES containing 15 nM pCellFreeG03+BMP2. Max fluorescence intensity (RFU) recorded over a 16 hours timelapse, where the blue colour corresponds to higher RFU values, and red to lower RFU values. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 1.

The highest values on the heatmap are attributed to the reactions containing 0.04 mM IPTG and no T7 RNAP. However, those values were a lot lower in comparison to the average 1 nM pdeGFP reaction at $15,098 \pm 2,216$ RFU (gain 70). In Figure 3.5 it was also difficult to determine any pattern associated with an increase or decrease in IPTG concentration across the reaction conditions tested here. Upon further examination and confirmation from the original pCellFree vector creator, Kirill Alexandrov, it was determined that the lac operon is a remnant of the vector backbone and no longer plays a role in protein expression¹⁵. Another relevant aspect for cell-free transcription is the presence of ribonucleases (RNases) which may contribute to the degradation of the

transcript, lowering the yield. In Figure 3.6, the addition of RNase inhibitors to the CFES reactions was summarised. RNase inhibitors are recombinant enzymes that inhibit the activity of the ribonucleases³³. Through the addition of these inhibitors, at least one aspect of influence on the protein production can be eliminated. Since TX-TL using pCellFree plasmids has not been successful with the in-house CFES, to determine whether these plasmids are indeed functional, PURExpress was also tested.

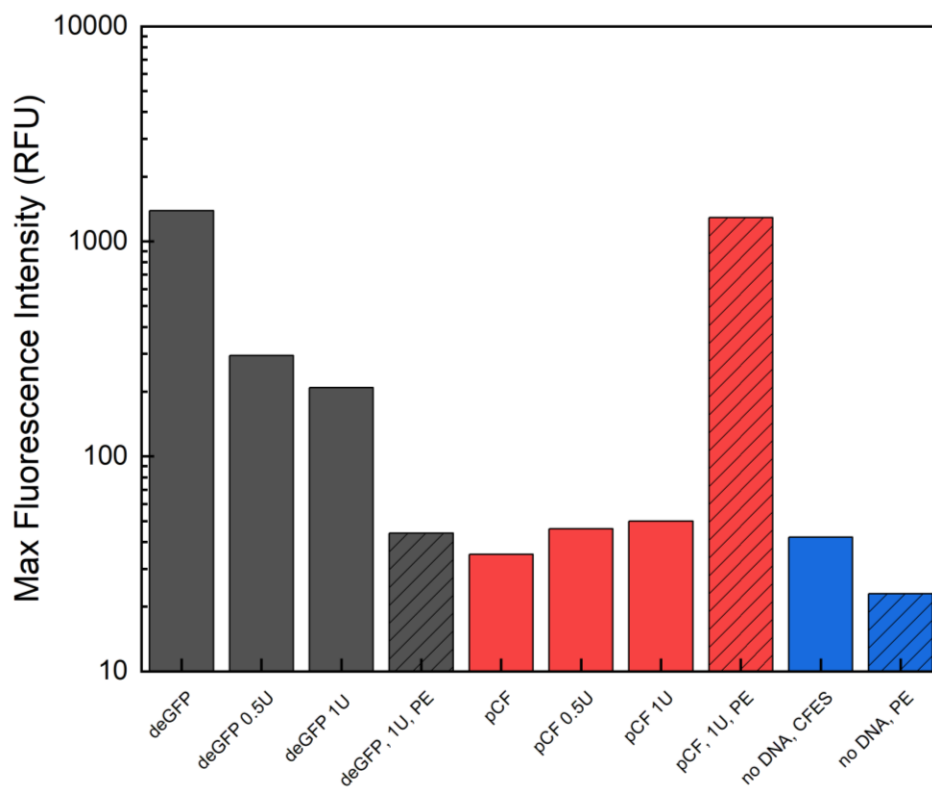


Figure 3.6 Maximum fluorescence intensity over a 16 hours time-lapse comparing pdeGFP and pCellFreeG03 expression from in-house or PURExpress CFES. 15 nM pdeGFP (black) and 15 nM pCellFreeG03 abbreviated as pCF (red) as DNA inputs for PURExpress (PE, diagonal lines) or in-house CFES (solid colour), with the addition of 4 U T7 RNAP (for in-house CFES) and 0.5 or 1 U of RNase inhibitor cocktail. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 50. Temperature= 37°C . Extract 2 for in-house CFES.

The highest recorded RFU in Figure 3.6 originated from pdeGFP containing reaction mixed with in-house CFES, as well as a reaction with pCellFreeG03 mixed with PURExpress. Increased concentration of RNase inhibitor added to the in-house CFES reaction produced a lower fluorescence signal of deGFP, on the contrary to pCellFreeG03.

The readings most similar to the background levels were that of pCellFreeG03 mixed with in-house CFES, and pdeGFP mixed with PURExpress. The most striking development depicted in Figure 3.6 is the positive fluorescence output when pCellFreeG03 was mixed with PURExpress, which suggests that this pCF plasmid is, in fact, functional. This result also implies that in-house CFES is not compatible with pCF plasmids, and PURExpress is not a suitable system for pdeGFP. PURExpress is generally the most high-yielding CFES, since it contains only the necessary recombinant machinery. However, the purity of the extract did not play the only crucial role here because PURExpress failed to warrant the expression of deGFP. The most obvious difference between the two systems is the type of RNA polymerase present in the mix. Potentially the reason why no expression was observed using in-house CFES supplemented with T7 RNAP was the unsuitable concentration, or perhaps the enzyme was inactive. Now that the pCellFreeG03 plasmid was transcribed and translated with PURExpress, the rest of the selected pCF plasmids, modified and unmodified, were tested with PURExpress (Figure 3.7).

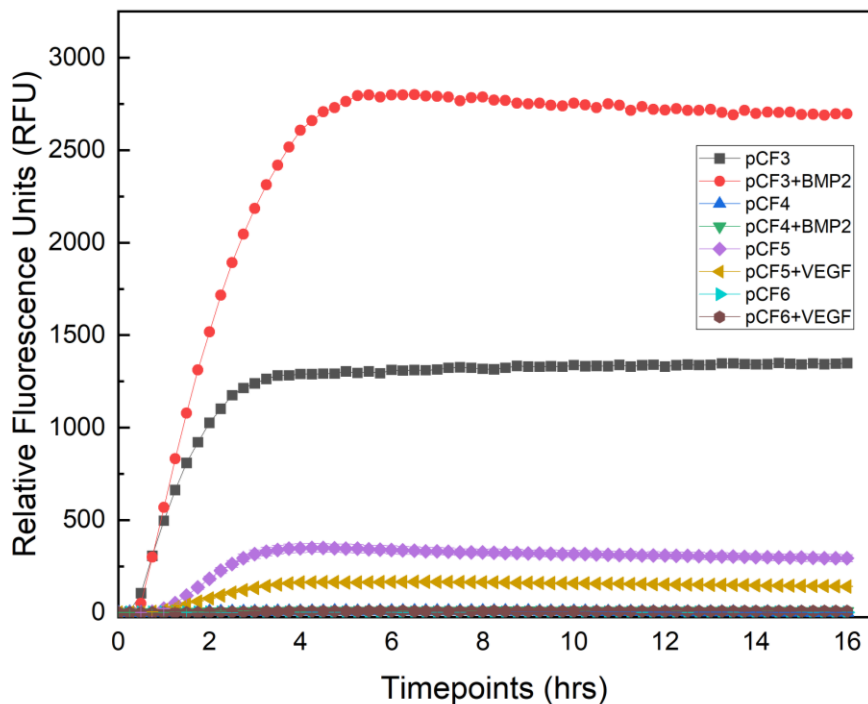


Figure 3.7 Kinetics of EGFP (pCF3 and pCF4), mCherry (pCF5 and pCF6) and their fusions with BMP2 or VEGF expression from 15 nM input plasmids with PURExpress. Background fluorescence subtracted. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm for EGFP variants. Excitation= 579 ± 10 nm, emission= 616 ± 20 nm for mCherry variants. Gain= 50. Temperature= 37°C .

The output of fluorescence over time is shown in Figure 3.7. In general, it follows the typical expression profile observed in cell-free TX-TL reactions, with an initial sharp linear increase within the first few hours followed by a plateau after about 5 hours, when the reaction is concluded. Here, similar RFU values from pCellFreeG03 reactions in PURExpress have been obtained as in Figure 3.6, showing a strong positive signal. Interestingly, a higher output was attained by pCellFreeG03+BMP2 of the same input concentration, which might suggest that by inserting the BMP2 sequence into this plasmid, the efficiency of TX-TL increased. From Figure 3.7, it is also possible to distinguish a positive signal from N-terminal mCherry-VEGF (pCellFreeG05+VEGF) and mCherry (pCellFreeG05), although the signal is relatively weak. It is not clear, however, why pCellFreeG06 and pCellFreeG06+VEGF did not bring about a response. Since it was shown in Figures 3.6 and 3.7 that some of the pCF plasmids can be functional, it was

important to investigate different cell-free systems to confirm the polymerase theory. To achieve this, a commercially available *E.coli* S30 extract based *in vitro* TX-TL kit was used, alongside in-house CFES and PURExpress. The S30 extract is supplemented with an optimized amount of T7 RNAP, but it also contains the endogenous polymerase in the extract. Therefore, in theory, the S30 commercial extract should be compatible with pdeGFP plasmid and the pCF plasmids. This is evidenced in Figure 3.8, where the highest fluorescence outputs for pCF plasmids are produced by PURExpress (diagonal lines) and the S30 commercial extract (solid colour). The fluorophore, deGFP, was successfully expressed by the S30 extract and the in-house CFES, but not PURExpress, as previously shown in Figure 3.6. Fluorescence output from the S30 reaction was higher than in-house CFES for deGFP expression, likely due to strict and repeatable extract protocols that are in place for its commercialisation. For pCF plasmids, the output from PURExpress was stronger than from S30, with the unusual exception of pCellFreeG04+BMP2. This could be attributed to the purity of PURExpress, or to the presence of only one polymerase, the T7 RNAP. The signal from in-house CFES with pCellFreeG05+VEGF and pCellFreeG06+VEGF was comparable to the background level and therefore cannot be seen on the graph in Figure 3.8. The important information that was decoded from this experiment was that deGFP and N-terminal EGFP-BMP2 (pCellFreeG03+BMP2) expression levels were comparable when suitable cell-free systems were chosen, and that the type of RNA polymerase present and its concentration plays a crucial part.

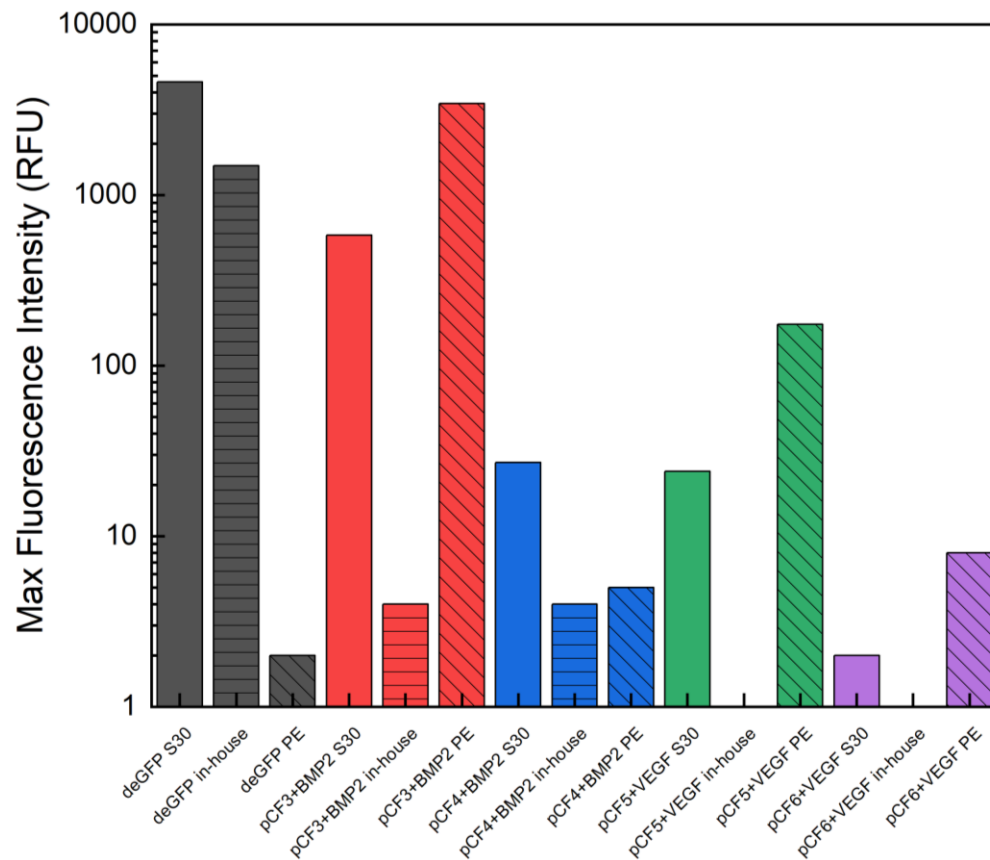


Figure 3.8 Maximum fluorescence intensity over a 16 hour time-lapse. Reactions of input DNA with three different sources of CFES, S30 extract-based T7 Promega kit (solid colour), in-house CFES (horizontal lines), and PURExpress (diagonal lines). 15 nM input DNA in the form of pdeGFP (black), pCF3+BMP2 (red), pCF4+BMP2 (blue), pCF5+VEGF (green) and pCF6+VEGF (purple). Background fluorescence subtracted. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm for EGFP variants. Excitation= 579 ± 10 nm, emission= 616 ± 20 nm for mCherry variants. Gain= 50. Temperature= 37°C . Extract 2 for in-house CFES.

To make in-house CFES a suitable environment for the transcription and translation of pCF plasmids, another method of supplementing T7 RNAP was investigated. Pr1-T7RNAP is a plasmid with OR2-OR1-Pr1 promoter expressing T7 RNAP developed by Siegal-Gaskins *et al.*³⁴. The presence of OR2-PR1-Pr1 promoter would allow for the recognition by the endogenous *E.coli* RNA polymerase that is contained within the in-house CFES. The transcript would be expressed into T7 RNAP, which in turn would recognise and bind to the T7 promoter on pCF plasmids, leading to the expression of

EGFP-BMP2 and mCherry-VEGF variants. This principle was tested with the in-house CFES along with the pCellFreeG03+BMP2 and presented in Figure 3.9.

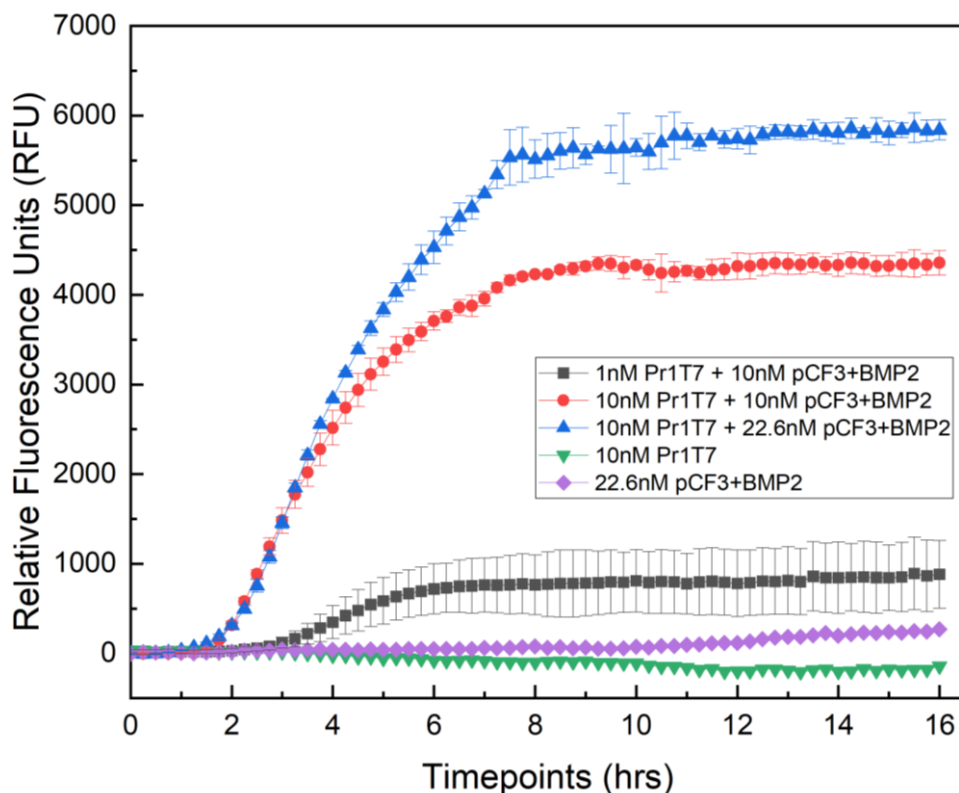


Figure 3.9 Kinetics of EGFP-BMP2 over a 16-hour timelapse, with Pr1T7 and pCF3+BMP2 at various concentrations as input DNA with in-house CFES. Background fluorescence subtracted. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 50. Temperature= 29°C . Extract 4. Error bars= standard deviation. Three technical repeats.

A clear increase in fluorescence was observed in the tested samples containing both plasmids in Figure 3.9. The only samples not presenting the sigmoidal curve shape were the negative reactions with singular plasmids. In the positive reactions, the increase in fluorescence was delayed to about 1.5 hours for 10 nM Pr1T7 with 10 nM or 22.6 nM pCF3+BMP2, and 3 hours for 1 nM Pr1T7 with 10 nM pCF3+BMP2, post reaction initiation. The delay is explained by the need to produce T7 RNAP from the Pr1T7 plasmid first before the expression from pCF3+BMP2. These two reactions were concluded by the 7 hour mark, whereas 1 nM Pr1T7 with 10 nM pCF3+BMP2 reaction finished sooner. The amount of the expressed fusion was manipulated by the

concentrations of the input DNA. Although the levels cannot be directly compared because of a different reaction temperature, it can be claimed that the 10 nM PrT7 with 10 nM pCF3+BMP2 TX-TL with the in-house CFES surpasses the amounts created from 15 nM pCF3+BMP2 TX-TL with PURExpress in Figure 3.8. These results prove the possibility of expressing EGFP-BMP2 from pCellFreeG03+BMP2 with the in-house CFES if Pr1T7 plasmid is also present.

3.3.3 TWO PLASMID SYSTEM

A two-plasmid system with pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 and pTar-AqpZ was designed specifically to produce and examine membrane proteins *in vitro*³⁵. Aquaporin Z (AqpZ) is an integral membrane protein responsible for osmoregulation in *E.coli*³⁶. Production of membrane proteins by CFES has now materialised as a great alternative tool, due to it being non-membrane bound, accessible and high-yielding³⁷. Here, both mentioned plasmids were donated by Dr Angelique Coutable (School of Biochemistry, University of Bristol). Prior to modifying the pTar-AqpZ plasmid to replace AqpZ gene with BMP2 sequence, the plasmids were tested with the in-house CFES (Figure 3.10). Fluorescence signal from deGFP is shown to surpass the signal from other fluorescent proteins in Figure 3.10. However, Figure 3.10 B shows that the two negative controls, 28 only and Aqpz-EGFP only, were both near zero fluorescence units. The σ 28 plasmid does not contain a fluorescent reporter gene, and so to achieve the expression of AqpZ-EGFP, the other plasmid is necessary. Considering this, no fluorescence output should be expected. There is a positive signal when the two plasmids were mixed in all four of the tested ratios. Out of the four, the lowest fluorescence signal is attained from 1 nM σ 28 with 1 nM AqpZ-EGFP in the in-house extract, which reached a plateau just below 300 RFU. There does not appear to be a significant difference in the intensity between the other three ratios. The initial rate of fluorescence increase occurs faster with these three ratios when compared to the 1 nM σ 28 with 1 nM AqpZ-EGFP, which can be explained by the amount of available initial DNA substrate.

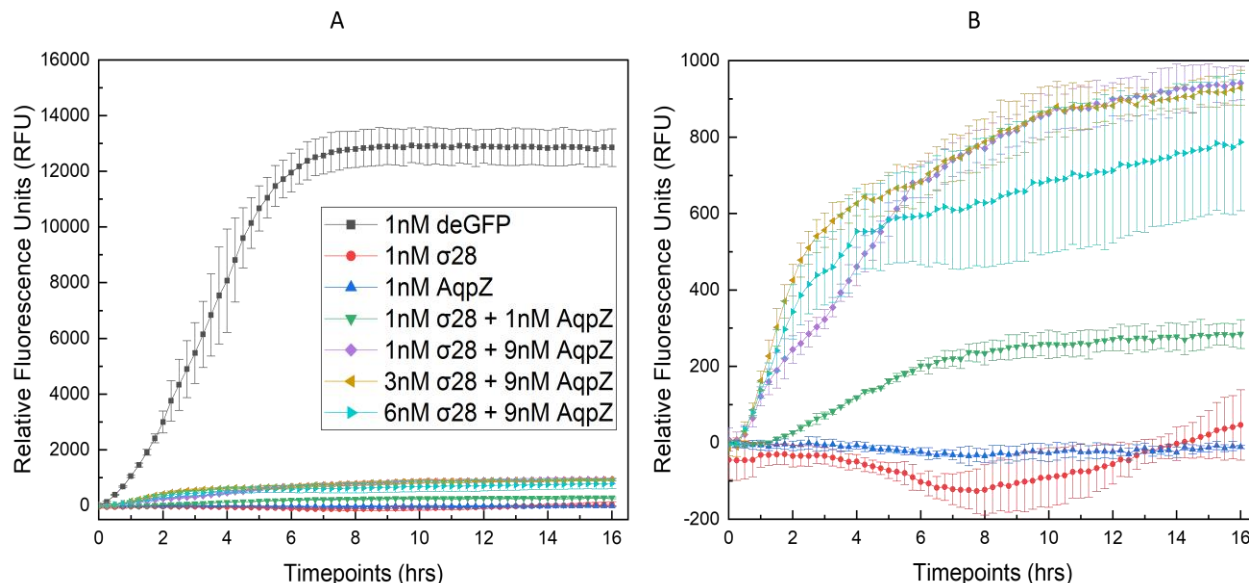


Figure 3.10 Kinetics of deGFP (black) AqpZ-EGFP (colour) over a 16-hour timelapse, with 1 nM pBEST-OR2-OR1-Pr-UTR1-deGFP-T500, and pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 with pTar-AqpZ at various concentrations as input DNA with in-house CFES. Background fluorescence subtracted. B: Rescaled graph with deGFP excluded. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C . Extract 1. Error bars= standard deviation. Three technical repeats.

Following this experiment, the AqpZ-EGFP plasmid was modified to replace the AqpZ gene with the BMP2 sequence. The extracted DNA from successful colonies post ligation was used to assess the plasmid's compatibility with the in-house extract along with the σ 28 plasmid, as illustrated in Figure 3.11. The two-plasmid system of σ 28 with BMP2-EGFP appeared to be functional with the in-house CFES, because of the positive increase of fluorescence over time.

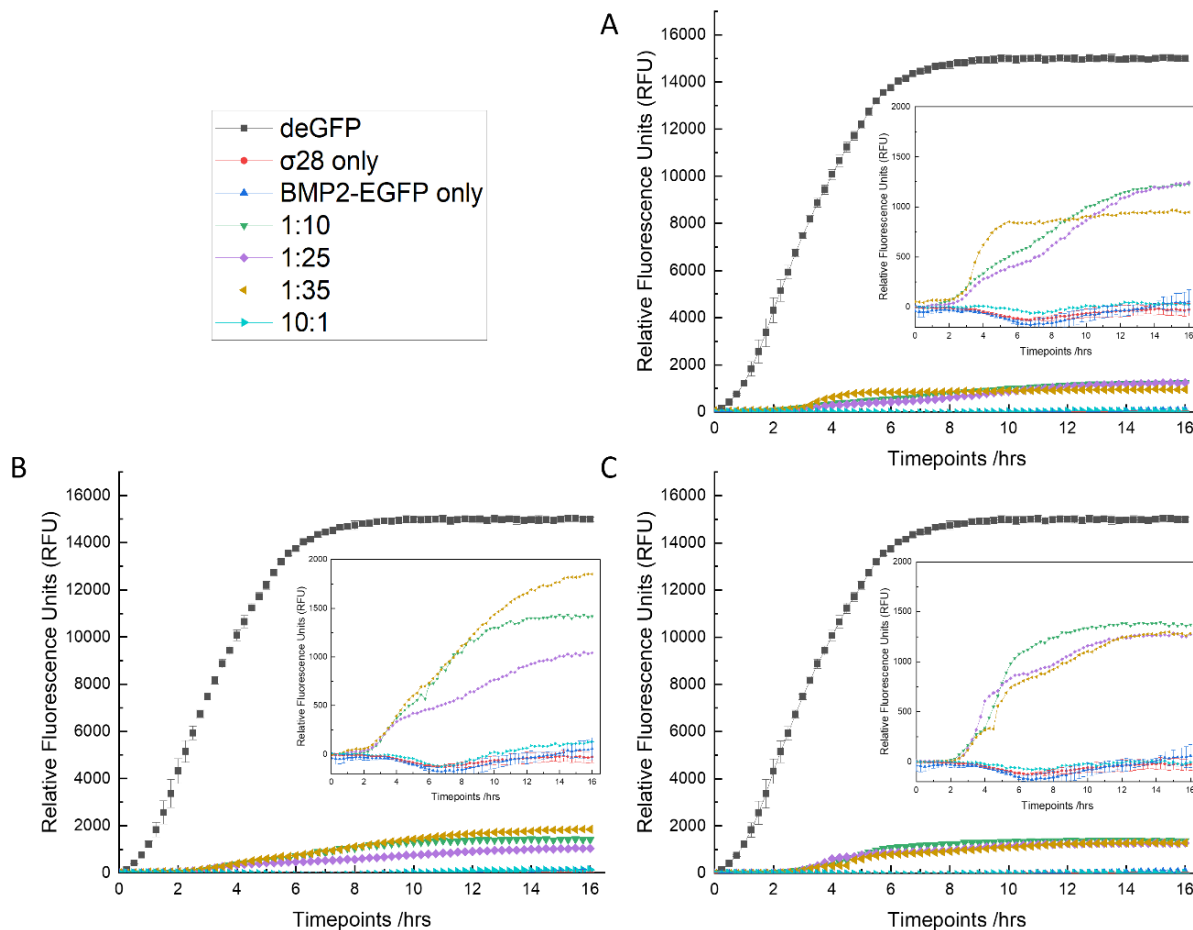


Figure 3.11 Kinetics of deGFP (black) and BMP2-EGFP (colours) over a 16-hour timelapse. DNA from three different colonies at various ratios with $\sigma 28$ plasmid. Inserts: deGFP excluded. A: DNA from colony numbered 2, B: DNA from colony numbered 5, C: DNA from colony numbered 8. Background fluorescence subtracted. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C . Extract 1.

The overall profile between the positive control, deGFP, and the samples differed in the initial reaction rate. deGFP can be detected early in the monitoring, whereas there is a time lag before BMP2-EGFP is expressed. This can be explained by the need to transcribe and translate $\sigma 28$ before BMP2-EGFP is made. Out of the four concentration ratios between the two plasmids tested, 10:1 (i.e. $8 \text{ nM } \sigma 28 : 0.8 \text{ nM BMP2-deGFP}$) is the only unsuccessful in expressing the protein. The rest of the tested ratios across colonies 2, 5 and 8 plateaued around 1000-2000 RFU, however a pattern could not be determined to find the most optimal ratio. The fluorescence levels of BMP2-EGFP are similar to the

levels of AqpZ-EGFP in the previous experiment (Figure 3.10), which plateaued just below 1000 RFU. It is likely that the expression yield would improve by investigating other ratios of the two plasmids.

3.3.4 DEGFP-BMP2

It has now been shown that pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 reference plasmid is superior to the other input DNA used with the in-house CFES. This conclusion resulted in the modification of this reference plasmid to include a growth factor, BMP2, sequence. Two variants were designed, where the BMP2 sequence was inserted either upstream or downstream from the deGFP open reading frame. Cell-free expression experiments using C-terminal deGFP-BMP2 as input DNA were ineffective, and the results are not shown here. On the other hand, N-terminal deGFP-BMP2 was probed further after displaying a positive result with in-house CFES. After successful ligation, several colonies were subject to DNA mini-prep and the plasmid concentration was determined, before being investigated with the in-house extract.

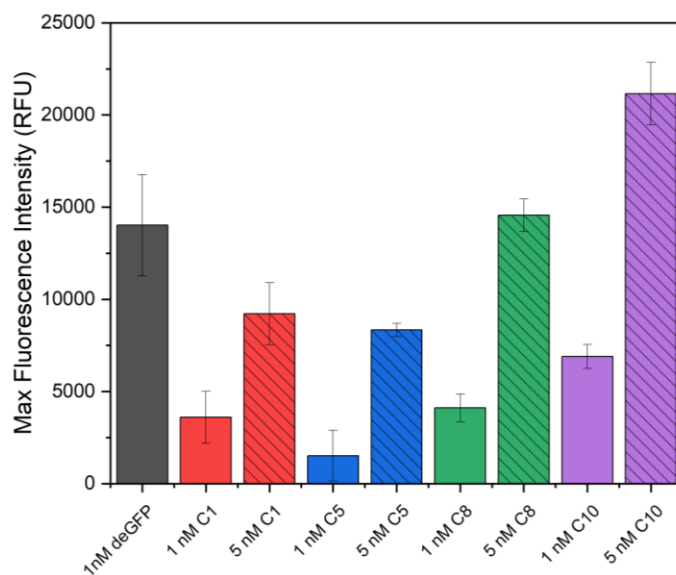


Figure 3.12 Maximum fluorescence intensity of deGFP-BMP2 over a 16-hour time-lapse. 1 nM and 5 nM input DNA originating from colonies numbered 1, 5, 8 and 10. Fluorescence from C2, C3, C4, C6, C7 and C9 colonies was similar to background levels and therefore not included. Background fluorescence subtracted. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C . Extract 1. Error bars= standard deviation. Three technical repeats.

When 1 nM input plasmids are compared in Figure 3.12, the strongest fluorescence signal was obtained from deGFP, although the fluorescence from the experimental samples have significantly improved when compared to the pCF plasmids with in-house extract in Chapter 3, Section 3.3.2, or the two-plasmid system in Chapter 3, Section 3.3.3. Significant fluorescence was detected from deGFP-BMP2 coming from all four represented colonies, with colony 10 DNA generating the strongest signal. When the DNA from C10 was increased to 5 nM, the amount of fluorescence output surpassed that of 1 nM pdeGFP, with maximum RFU at 21155. This high fluorescence signal indicated a high yield of deGFP-BMP2 protein and demonstrated the compatibility of this plasmid with the in-house extract. The molecular manipulation of the plasmid did not impede the expression of deGFP significantly, and one method of increasing the output could be performed by regulating the input DNA concentration, which was studied further in Figure 3.13.

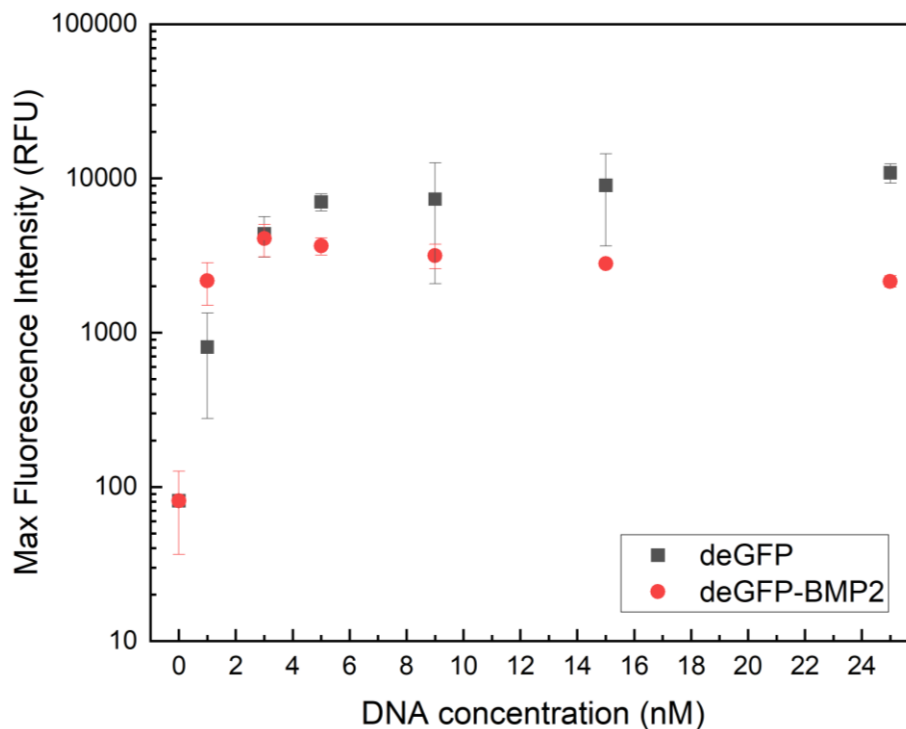


Figure 3.13 Maximum fluorescence intensity of deGFP and deGFP-BMP2 over a 16-hour time-lapse, with various input DNA concentrations. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 50. Temperature= 29°C . Extract 4. Error bars= standard deviation. Three technical repeats.

Figure 3.13 displays the dependence of input DNA concentration on the fluorescent chimera production with the in-house CFES. The signal at 0 nM DNA indicates the fluorescence background of the extract. When 1 nM plasmids are utilized, the output of the chimera is x2.5 higher in comparison to deGFP, but at 3 nM DNA concentration, the outputs of both proteins are similar. However, at higher concentrations, the signal from deGFP is above that of the chimera. There is no significant increase in the signal from both protein when DNA concentration was elevated beyond 5 nM in Figure 3.13. This is commonly observed in literature, where a threshold is reached, not producing larger amounts of protein at higher DNA concentrations³⁸. Although there are numerous speculations as to which factors are rate-limiting, discussed further in Chapter 1, Section 1.2, this threshold can be associated with the saturation in the translation machinery¹¹. Unless those rate-limiting factors are addressed, the threshold will not be mitigated. Here, multitudinous factors, such as NTPs, elongation factors and many more were not investigated, nonetheless, the DNA concentration dependency experiment allowed to identify the most efficient input plasmid concentration, using the minimal amount of substrates.

Performing CFES reactions near the physiological temperature is an important aspect to consider if the system is to be adapted further into an injectable hydrogel or if spatiotemporal protein synthesis is to be controlled by external factors such as temperature. The deGFP-BMP2 miniprep from Colony 10 was shown to be a reliable, high-yielding growth factor fusion which was chosen for further experiments. Therefore, the temperature effect on CFES was measured with the fusion as input DNA as seen in Figure 3.14. The fluorescence levels of the chimera and the pattern across the DNA concentration at 29°C are observed to be same in Figure 3.13 and 3.14. In Figure 3.14, the pattern of fluorescence output dependency on DNA concentration is the same at both reaction temperatures, however, the signal is significantly reduced when the reaction is incubated at 37°C. The lower reaction temperature could be enhancing the protein folding

dynamics, leading to a higher soluble protein yield³⁹. Nonetheless, the significant positive signal at 37°C, even at DNA concentrations as low as 1 nM was satisfactory to confirm chimera production at the physiologically-relevant temperature.

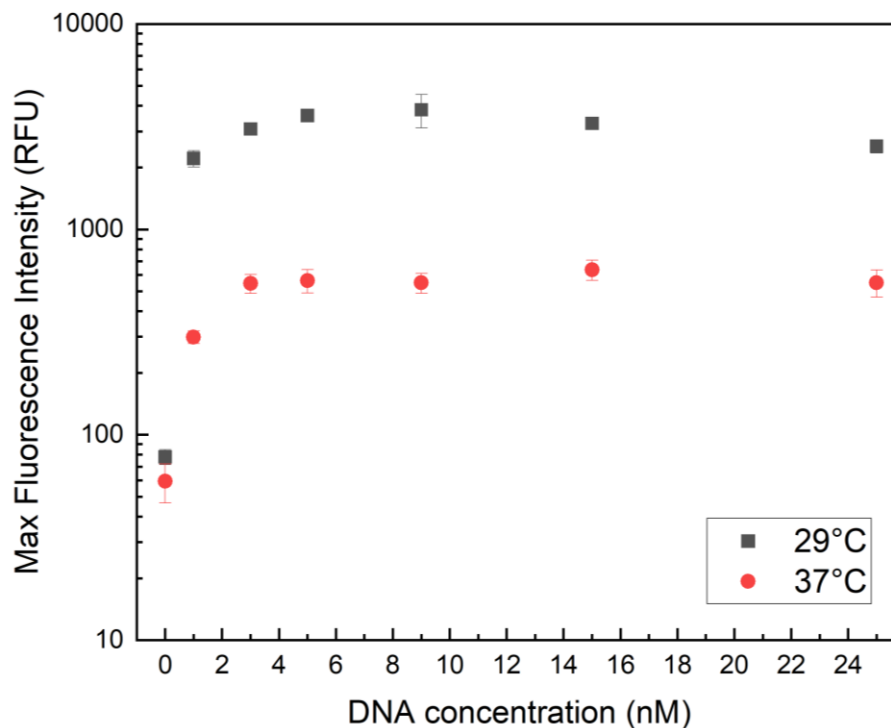


Figure 3.14 Maximum fluorescence intensity of deGFP-BMP2 over a 16-hour time-lapse, with the CFES reaction containing various input DNA concentrations, and carried out at 29°C and 37°C. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 50. Extract 4. Error bars= standard deviation. Three technical repeats.

Considering that the strongest fluorescence output of the reporter protein-growth factor fusion was obtained from deGFP-BMP2, out of all the designed plasmids, it was selected for further experimentation. It was important to select the highest yielding conditions, and therefore, after some observations, the extracts were directly compared to produce the chimera. The three extract reactions depicted in Figure 3.15, show a similar transient phase for the first 30 minutes to 1 hour, followed by a steady state of gene expression, before the reactions expire roughly 8 hours post-initiation. The highest production of the chimera was obtained from CFES when extract 4 was used. Out of the three tested

variants, extract 3 produced the lowest yield of the protein. This can be attributed to the quality of the extract made, which is discussed in Section 3.2.1.

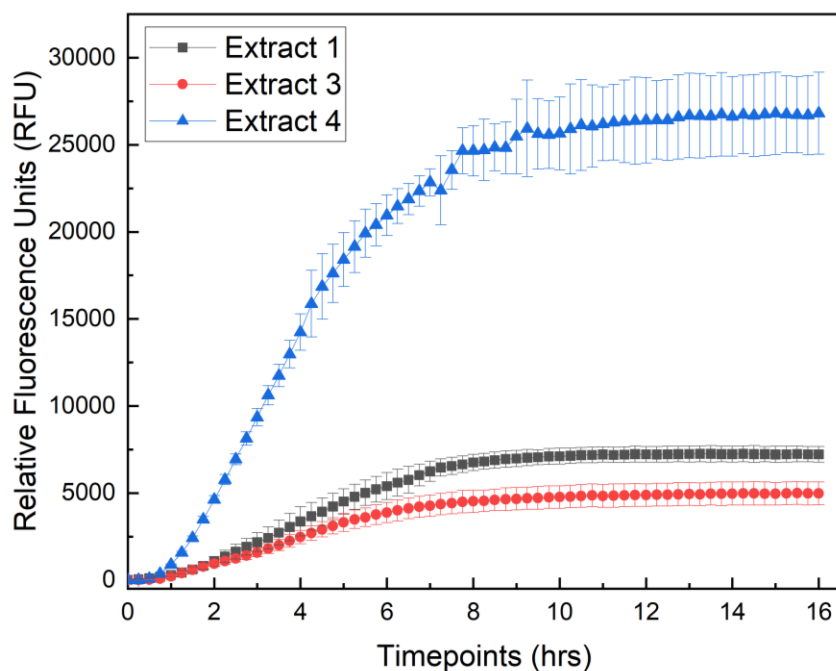


Figure 3.15 Kinetics of deGFP-BMP2 chimera produced from three extracts over a 16-hour timelapse. 1 nM deGFP-BMP2 as input DNA. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C . Error bars= standard deviation. Three technical repeats.

The differentiation assay experiments involving CFES (Chapter 5) were performed with extract 4 and pdeGFP-BMP2. To determine the concentration of the CFES produced deGFP-BMP2 using fluorescence outputs, a standard curve was created. The standard curve was produced using extract 4 supplemented with a concentration range of *E.coli* expressed deGFP-BMP2 (Figure 3.16).

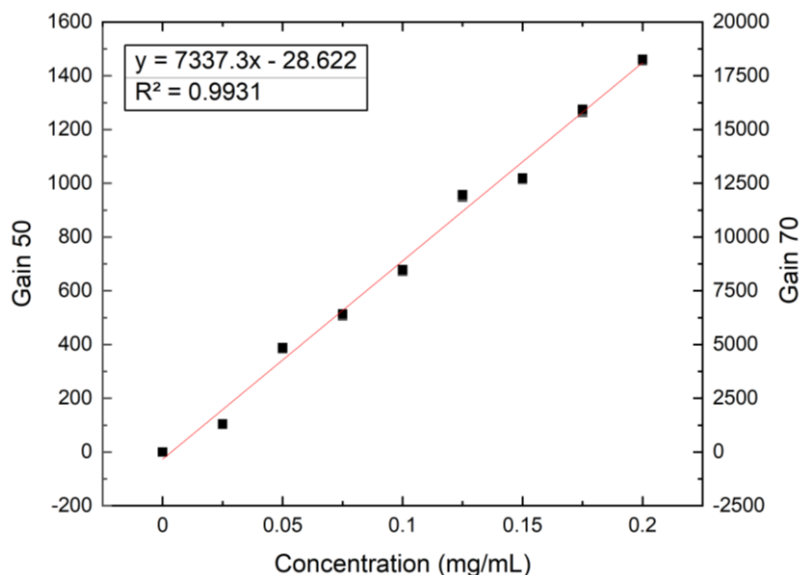


Figure 3.16 Standard curve of deGFP-BMP2 in extract 4 CFES. Measured as relative fluorescence units at gain of 50 and gain of 70 of the plate reader. Temperature= 29°C. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Three technical repeats.

Based on the standard curves including the one presented in Figure 3.16, a summary of the deGFP-BMP2 average yields from CFES reactions were assembled into Table 3.1. The highest obtainable yield was measured at 0.49 mg/mL, whereas unmodified deGFP collected by Shin *et al.*¹⁰ was quoted at 0.65 mg/mL.

DNA conc. (nM)	Yield (mg/mL)	(μ M)
1	0.30	7.49
3	0.49	12.21
5	0.49	12.36
9	0.48	11.94
15	0.42	10.38
25	0.32	8.02

Table 3.1 Average concentrations of deGFP-BMP2 from in-house CFES reactions carried out at 29°C with different input DNA concentrations.

The reduction of yield between deGFP and the chimera, observable in Figure 3.13, can be associated with the lower fluorescence level for the chimera. The regulatory elements

present on pBEST-OR2-OR1-Pr-UTR1-deGFP-T500 and on the chimera plasmid were unmodified nor deleted; as such, the output difference was likely linked to the addition of the BMP2 sequence. deGFP-BMP2 is a larger product in comparison to deGFP, meaning more substrates and energy would be required for its production, leading to faster deficiency, and lowering of the system's efficiency. However, in a publication by Gagoski *et al.*, a synthesised protein's size did not show a significant correlation to the obtained yield in *E.coli* cell-free systems⁴⁰. In fact, Gagoski *et al.* suggested that protein TX-TL finishes before the substrates are spent, and propounding that a change in the pH or a change in the crowding within the reaction could be a contributing factor for lower yields generated by larger plasmids.

3.3.5 CELL-FREE PROTEIN SYNTHESIS IN HYDROGELS

Generally, cell-free protein synthesis is completed in the liquid phase. However, in recent years, there has been an explosion of interest in carrying out CFES in various chassis, where this variety was discussed in Chapter 1, Section 1.1.7. Here, some preliminary experiments were performed with CFES in hydrogels, including agar, agarose, and an F127-alginate bioink. One of the first experiments consisted of mixing pdeGFP CFES reactions with separate bioink components (Figure 3.17) to observe whether protein synthesis is possible within alginate and F127. Figure 3.17 A illustrates a positive fluorescence signal from all of the pdeGFP CFES containing reactions. The transient phase is observed for a much longer time period than in 10 μ L liquid CFES. The 2-fold dilution of the CFES reaction with either water or the alginate/F127 contributed to the minimization of the macromolecular crowding effect and the dilution lowered the possible biomolecule interactions⁴¹. Due to the prolonged transient phase, the reactions did not reach a plateau after 23 hours. However, the signal trends of the gel reactions were similar to the liquid 20 μ L CFES (black, square), especially 1% w/v alginate CFES (green, triangle). The error bars are large, with an extensive variation of the readings from sample repeats.

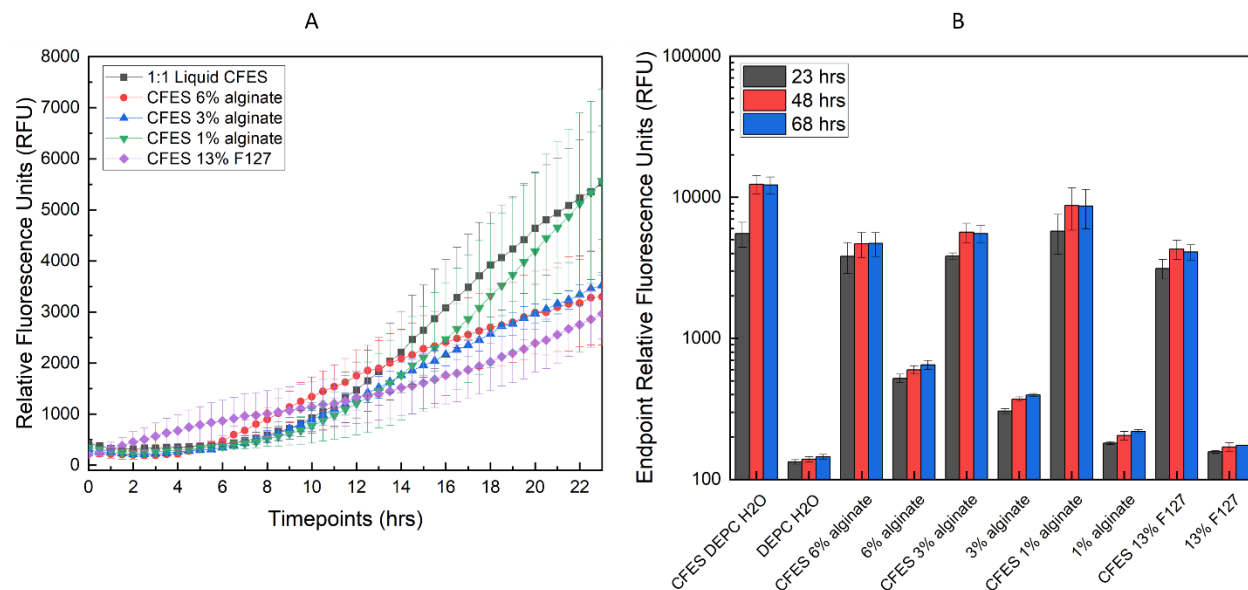


Figure 3.17 Kinetics of deGFP produced from bioink components containing CFES. A: 20 μL reactions consisting of 10 μL CFES with 9 nM input DNA and 10 μL sodium alginate at a final % w/v of 6, 3, 1 (red, blue, green respectively), or 10 μL DEPC water (black). 6.5 μL Pluronic F127 at a final 13 % w/v with 13.5 μL CFES with 9 nM input DNA (purple). Background fluorescence subtracted. B: Endpoint fluorescence readings 23 hours (black), 48 hours (red) and 68 hours (blue) post-initiation of reactions from A. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 90. Temperature= 29°C . Extract 2. Error bars= standard deviation. Three technical repeats.

The fluorescence readings from the same samples were then measured 48 hours (red columns) and 68 hours (blue columns) post initiation (Figure 3.17 B) to determine reaction plateau. Mean endpoint RFU increased between 23 hrs and 48 hrs post reaction initiation. This suggested that the reaction endpoint was after 23 hrs but before 48 hrs, since the mean RFU at 68 hrs was consistent with 48 hrs RFU. This significant increase between 23 hrs and 48 hrs was observed in 3% w/v alginate CFES, 13% F127 CFES and the liquid CFES. It is worth noting that the lower the final % w/v of sodium alginate with CFES, the higher the protein expression of deGFP. The bioink components were then mixed as 2% w/v for 1:1 volume ratio (initial sodium alginate as 12% w/v, and F127 as 26% w/v), and 4% w/v for 1:3 volume ratio (initial sodium alginate as 24% w/v, and F127 as 52% w/v), and used for incorporation with CFES reactions (Figure 3.18).

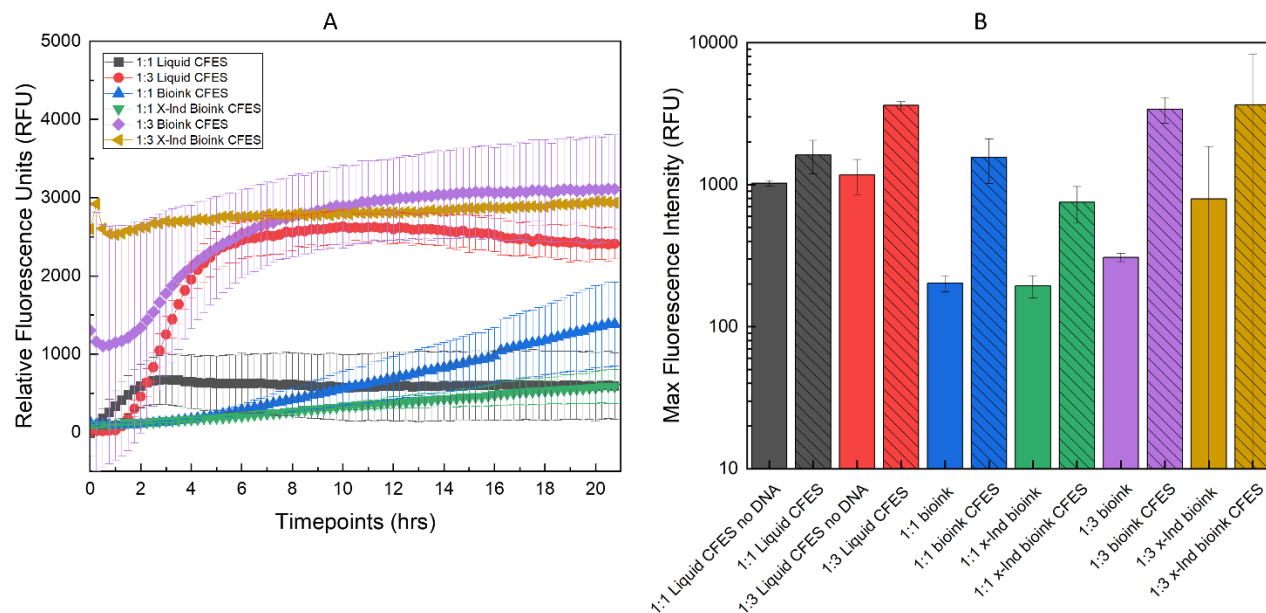


Figure 3.18 Kinetics of deGFP produced from bioink containing CFES, over a 21-hour time-lapse. A: 20 μ L reactions in 384 well plate, carried out as 1:1 volume ratios or 1:3 volume ratios with H₂O or CFES. 1:1 mixtures consist of 10 μ L DECP water (black, square)/bioink (blue, triangle)/cross-linked bioink (green, triangle) and 10 μ L CFES with 9 nM deGFP input DNA. 1:3 mixtures consist of 5 μ L DEPC water (red, circle)/bioink (purple, diamond)/cross-linked bioink (yellow, triangle) and 15 μ L CFES with 9 nM deGFP input DNA. Large 1:3 x-linked bioink error bars removed for clarity. Background fluorescence subtracted. B: Max fluorescence intensity representation of A as a bar chart, with negative controls (solid colour) and their CFES containing counterparts (diagonal lines). Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C . Extract 4. Error bars= standard deviation. Three technical repeats.

In Figure 3.18, the yield from liquid CFES is surprisingly low, and although there is a clear sigmoidal shape, the reaction expires exceptionally fast at around 3 hours post-initiation for 1:1 liquid CFES, and at around 6 hours post-initiation for 1:3 liquid CFES. Nonetheless, the difference between the positive control and the background is noted. The final composition of the bioink as 6% w/v sodium alginate and 13% w/v Pluronic F127 with CFES is represented as blue triangles (1:1) and purple diamonds (1:3). The 1:1 bioink CFES shows an increase in fluorescence, 5 hours post-initiation, but the maximum fluorescence intensity is found to be similar to the positive 1:1 liquid CFES. This is also the case for 1:3 bioink CFES, having a similar maximum fluorescence intensity to the 1:3 liquid CFES. Crosslinked bioink reactions were difficult to set up, and this translated to

very large errors (removed for 1:3 x-linked bioink CFES as this was spanning the whole graph view, Figure 3.18 A). Once the bioink is cross-linked with calcium chloride, it becomes opaque and therefore, challenging to obtain stable fluorescence readings from.

Mixing of pdeGFP CFES with agar and agarose was also investigated in a similar manner (Figure 3.19). It was found that the measured signal is much stronger for pdeGFP CFES with agar and agarose, in comparison to the bioink CFES (Figure 3.18), as well as bioink components, sodium alginate and Pluronic F127 (Figure 3.17).

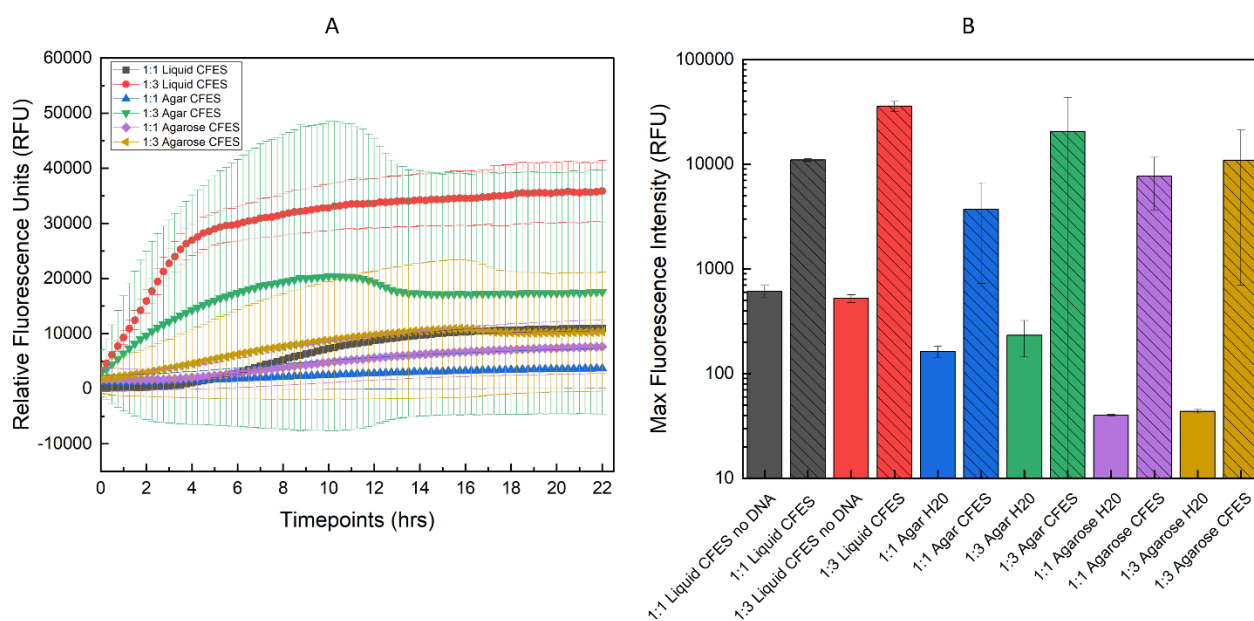


Figure 3.19 Kinetics of deGFP produced from agar and agarose containing CFES, over a 22-hour time-lapse. A: 20 μ L reactions in 384 well plate, carried out as 1:1 volume ratios or 1:3 volume ratios with H₂O or CFES. 1:1 mixtures consist of 10 μ L DECP water (black, square)/agar (blue, triangle)/agarose (purple, diamond) and 10 μ L CFES with 9 nM deGFP input DNA. 1:3 mixtures consist of 5 μ L DEPC water (red, circle)/agar (green, triangle)/agarose (yellow, triangle) and 15 μ L CFES with 9 nM deGFP input DNA Both agar and agarose had a final 1% w/v. Background fluorescence subtracted. B: Max fluorescence intensity representation of A as a bar chart, with negative controls (solid colour) and their CFES containing counterparts (diagonal lines). Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 4. Error bars= standard deviation. Three technical repeats.

The samples tested appear to have run to completion due to the presence of plateaus after 22 hours. Some of the error bars, especially from 1:3 agar CFES reaction repeats, are very large, which can be attributed to the ineffective method of mixing between the gels and

CFES. 1:3 liquid CFES produced more deGFP than 1:1 liquid CFES, due to it containing a higher volume percentage of extract and buffer. This difference is more difficult to distinguish with agar or agarose containing reactions due to the large errors, although the mean fluorescence intensity for 1:3 reactions are higher. Overall, there is a strong fluorescence signal from agar and agarose hydrogel CFES hybrids.

CFES hydrogels were further investigated, by separately monitoring the fluorescence from the gel area and the liquid surrounding it. By generating separate readings, the potential diffusion of the produced deGFP-BMP2 from the hydrogel CFES could be observed. This was conducted by pipetting the hydrogel CFES into a specific location of a well of a 96-well-plate, allowing time for gelation, followed by submersion in DEPC water (Figure 3.20). The fluorescence reading area was divided equally into nine regions, including the gel area (Position 1 of Figure 3.20) and the surrounds (Position 2 of Figure 3.20). Over a 20 hour time period, the fluorescence from the coordinates of the gels' location (Position 1) and the coordinates of the furthest area away from the gel (Position 2), are plotted in Figure 3.21 and Figure 3.22 respectively.

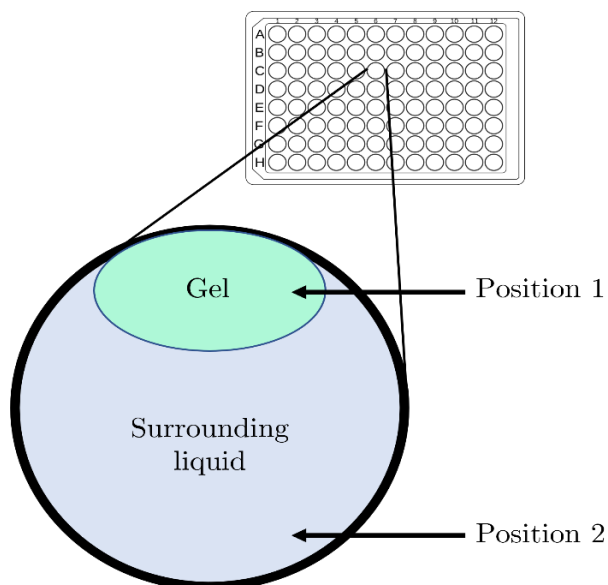


Figure 3.20 Diffusion monitoring set-up in a 96-well plate where the gel was pipetted to one side of the well and its fluorescence was measured at Position 1, and once the gel was set it was submerged in DEPC water where fluorescence was measured at Position 2.

The highest intensity noted from both figures comes from liquid CFES reactions, 1:1 and 1:3 v/v dilutions with DEPC water (wells 2 and 4, respectively). Fluorescence from 1:3 v/v diluted liquid CFES reaction (well 4) is logically higher than fluorescence from 1:1 v/v diluted liquid CFES (well 2) as the reaction contains a higher proportion of CFES components. This difference extends to the agar CFES gel (wells 6 and 8) but, in agarose the signal from 1:1 v/v CFES gel (well 10) is similar to the output from 1:3 v/v CFES gel (well 12) (Figure 3.21). The 1:3 diluted CFES-containing agar gel (well 8) had the same fluorescence as 1:3 agar gel with water (well 7) due to a technical error of adding an arbitrary amount of CFES reaction instead of water. A difference in the trendline shape and fluorescence amount can be perceived between a 1:1 v/v CFES agar (well 6) and 1:1 v/v water agar (well 5). This is the same as the agarose equivalent (wells 10 and 9 respectively), which could be attributed to the production of deGFP-BMP2 or just the presence of the auto-fluorescent CFES machinery. Unfortunately, without adequate controls, it is not possible to determine the source of the fluorescence signal. When the liquid CFES reactions were compared between Figure 3.21 and Figure 3.22, the fluorescence outputs are inconsistent between the two areas, whereas, because of their liquid property, the reactions should have dispersed over the well. The small reaction volume in a large well of a 96 well plate could perhaps not fill the whole area or results in a rapid evaporation. There is a lower fluorescence signal obtained from CFES-containing reactions (wells 6, 8, 10 and 12) in the surrounding water (Figure 3.22), in comparison to the gel location (Figure 3.21), however, the difference between CFES +ve and CFES -ve reactions is still observed at both locations. Any increase in the signal in Figure 3.22 could indicate some degree of leaching from the gel into the water surrounding it. At the end of the time-lapse, it was also noted that a significant portion of the liquid reactions have evaporated, which could be a source of further errors.

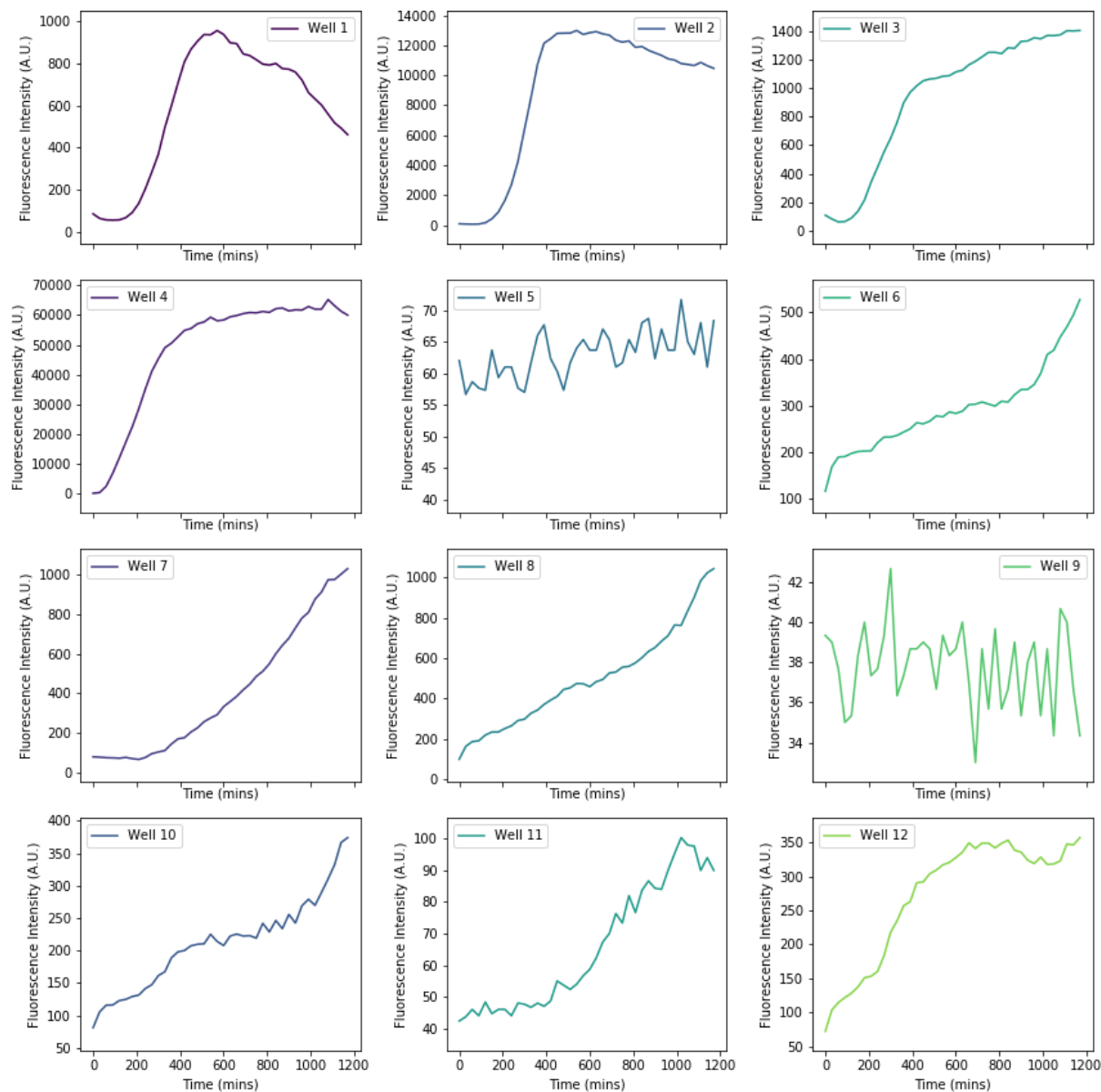


Figure 3.21 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area where the agar or agarose gels were placed in a 96-well plate (Position 1). Average of 3 technical repeats. Well 1 refers to a 1:1 diluted negative CFES (-DNA) with DEPC water. Well 2: a 1:1 diluted positive CFES (+DNA) with DEPC water. Well 3: a 1:3 diluted negative CFES (-DNA) with DEPC water. Well 4: a 1:3 diluted positive CFES (+DNA) with DEPC water. Well 5: a 1:1 v/v of agar (final 1% w/v) with DEPC water. Well 6: a 1:1 v/v of agar (final 1% w/v) with +ve CFES. Well 7: a 1:3 v/v of agar (final 1% w/v) with DEPC water. Well 8: a 1:3 v/v of agar (final 1% w/v) with +ve CFES. Well 9: a 1:1 v/v of agarose (final 1% w/v) with DEPC water. Well 10: a 1:1 v/v of agarose (final 1% w/v) with +ve CFES. Well 11: a 1:3 v/v of agarose (final 1% w/v) with DEPC water. Well 12: a 1:3 v/v of agarose (final 1% w/v) with +ve CFES. 1 nM pdeGFP-BMP2 as input DNA. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 4.

Chapter 3 - Optimisation of the Cell-Free Expression System

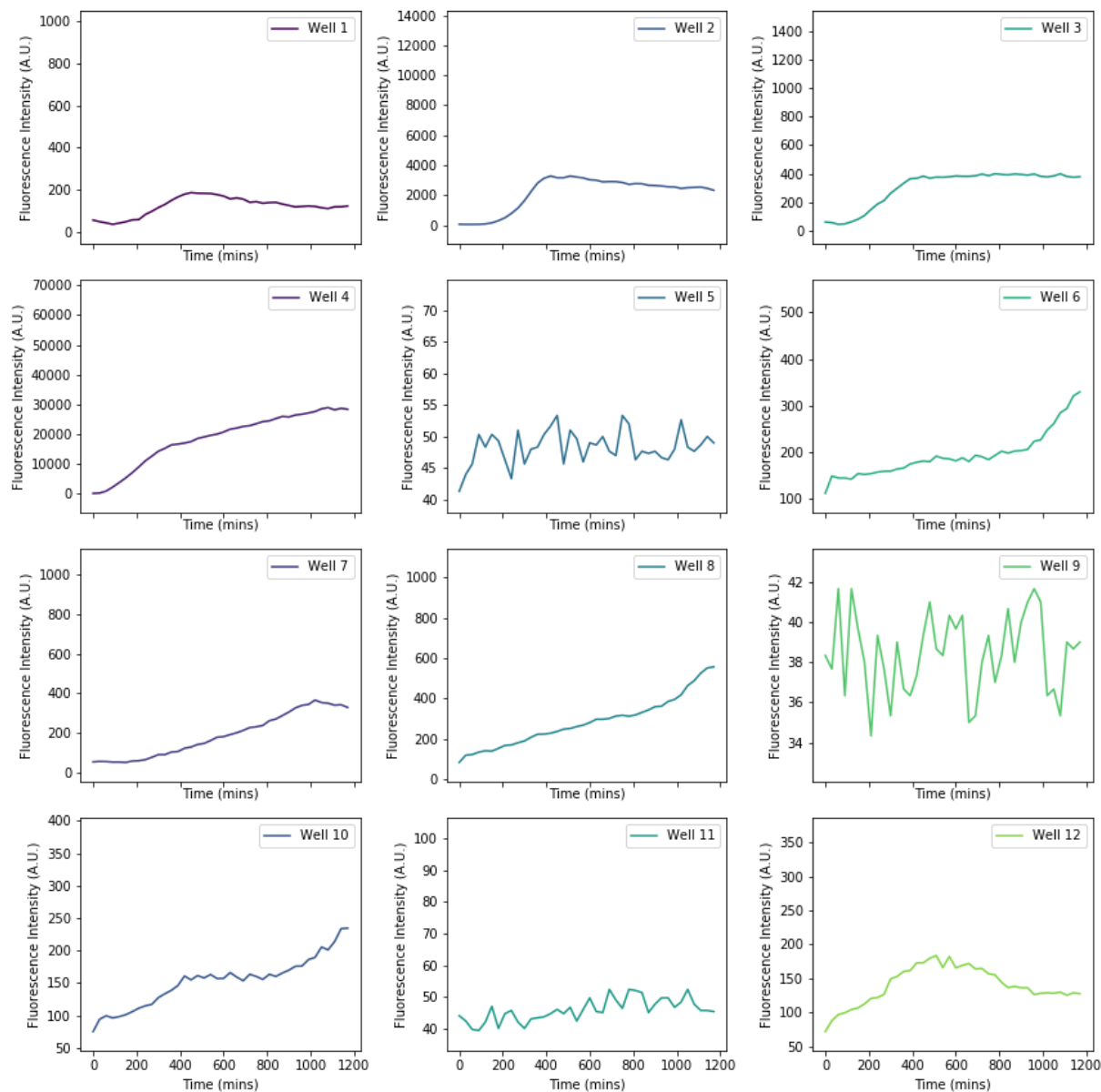


Figure 3.22 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area furthest away from where the agar or agarose gels were placed in a 96-well plate (Position 2). Average of 3 technical repeats. Well 1 refers to a 1:1 diluted negative CFES (-DNA) with DEPC water. Well 2: a 1:1 diluted positive CFES (+DNA) with DEPC water. Well 3: a 1:3 diluted negative CFES (-DNA) with DEPC water. Well 4: a 1:3 diluted positive CFES (+DNA) with DEPC water. Well 5: a 1:1 v/v of agar (final 1% w/v) with DEPC water. Well 6: a 1:1 v/v of agar (final 1% w/v) with +ve CFES. Well 7: a 1:3 v/v of agar (final 1% w/v) with DEPC water. Well 8: a 1:3 v/v of agar (final 1% w/v) with +ve CFES. Well 9: a 1:1 v/v of agarose (final 1% w/v) with DEPC water. Well 10: a 1:1 v/v of agarose (final 1% w/v) with +ve CFES. Well 11: a 1:3 v/v of agarose (final 1% w/v) with DEPC water. Well 12: a 1:3 v/v of agarose (final 1% w/v) with +ve CFES. 1 nM pdeGFP-BMP2 as input DNA. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 4.

The same setup was used to generate readings from bioink CFES (Figures 3.23 and 3.24), although, additional controls were put in place, including negative controls of bioink mixed with -ve CFES (no input DNA). The bioink gels were crosslinked with 100 mM CaCl₂ to ensure that the surrounding water did not dissolve it. In Figure 3.23, clear signal level differences between liquid CFES containing DNA or water, are observed as in previous Figures, indicating the production of deGFP-BMP2. The difference between +DNA (wells 6 and 9)/- DNA (wells 7 and 10) CFES with bioink is quite minimal and a clearer difference can be noted with bioink gels with no CFES added to them. When analysing Figure 3.24, it is seen that the water surrounding +DNA CFES bioink produces a stronger fluorescence signal than the water surrounding -DNA CFES bioink, which could be indicative of a diffusion of deGFP-BMP2 from the gels, although the differences are very low. In both cases of the agar/agarose and bioink diffusion monitoring experiments, the location of the deposited gels is largely estimated, therefore, to produce more accurate readings, a more suitable method of locating the gel should be undertaken, such as using grided wells. With low fluorescence output between -/+DNA CFES gels, it was likely the TX-TL machinery was contributing mostly to that signal. Sampling the surrounding water from the reactions, and dissolving the gels to produce a Western blot, would enable to distinguish the fluorescence contributions between deGFP-BMP2 and the CFES components.

Chapter 3 - Optimisation of the Cell-Free Expression System

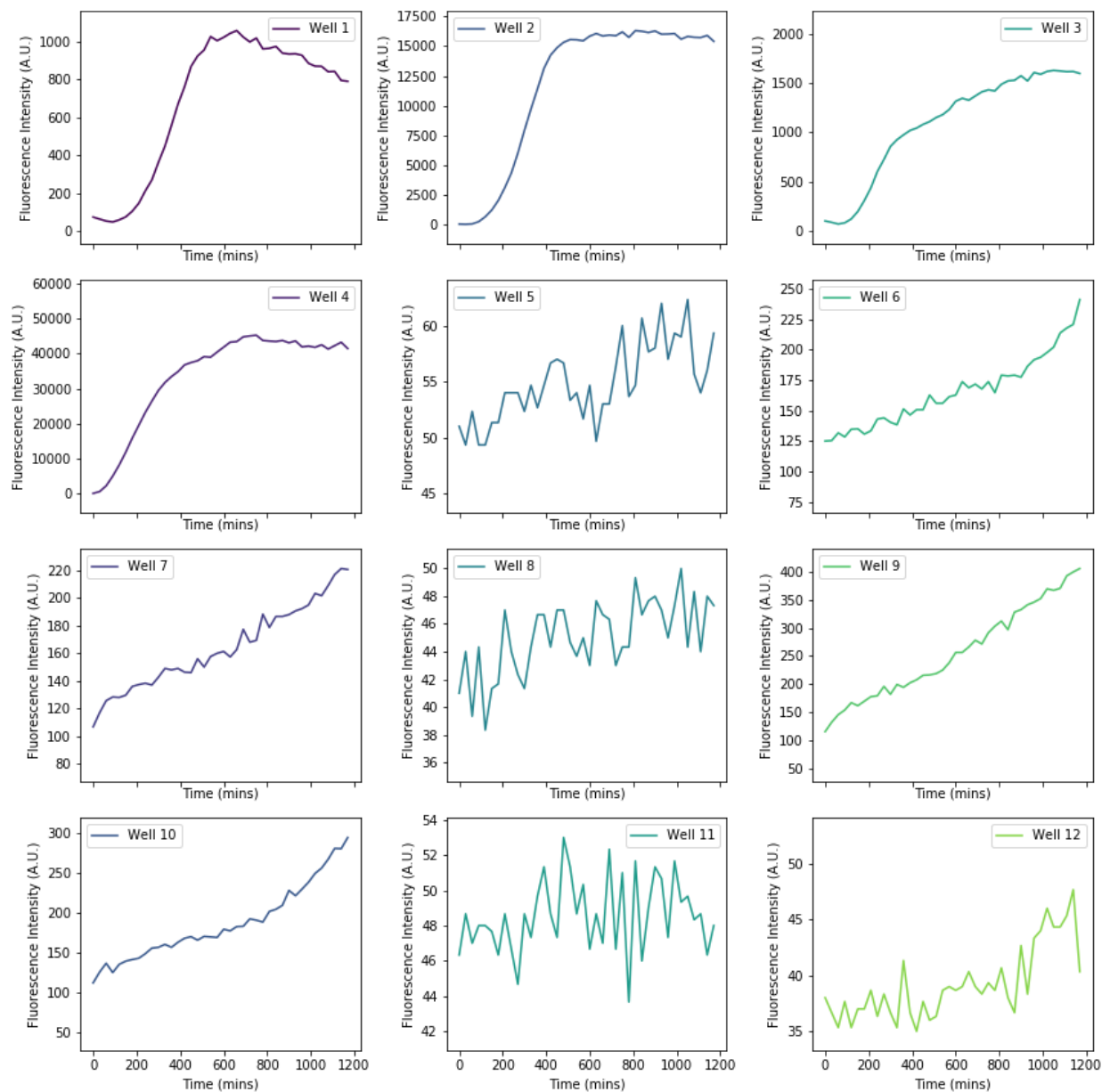


Figure 3.23 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area where the bioink gels were placed in a 96-well plate (Position 1). Average of 3 technical repeats. Well 1 refers to a 1:1 diluted negative CFES (-DNA) with DEPC water. Well 2: a 1:1 diluted positive CFES (+DNA) with DEPC water. Well 3: a 1:3 diluted negative CFES (-DNA) with DEPC water. Well 4: a 1:3 diluted positive CFES (+DNA) with DEPC water. Well 5: a 1:1 v/v of bioink (final 6% w/v alginate and 13% w/v F127) with DEPC water. Well 6: a 1:1 v/v of bioink with +ve CFES in 100 mM CaCl₂. Well 7: a 1:1 v/v of bioink with -ve CFES in 100 mM CaCl₂. Well 8: a 1:3 v/v of bioink (final 6% w/v alginate and 13% w/v F127) with DEPC water. Well 9: a 1:3 v/v of bioink with +ve CFES. Well 10: a 1:3 v/v of bioink with -ve CFES. Wells 11 and 12: empty. 1 nM pdeGFP-BMP2 as input DNA. Excitation= 479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 4.

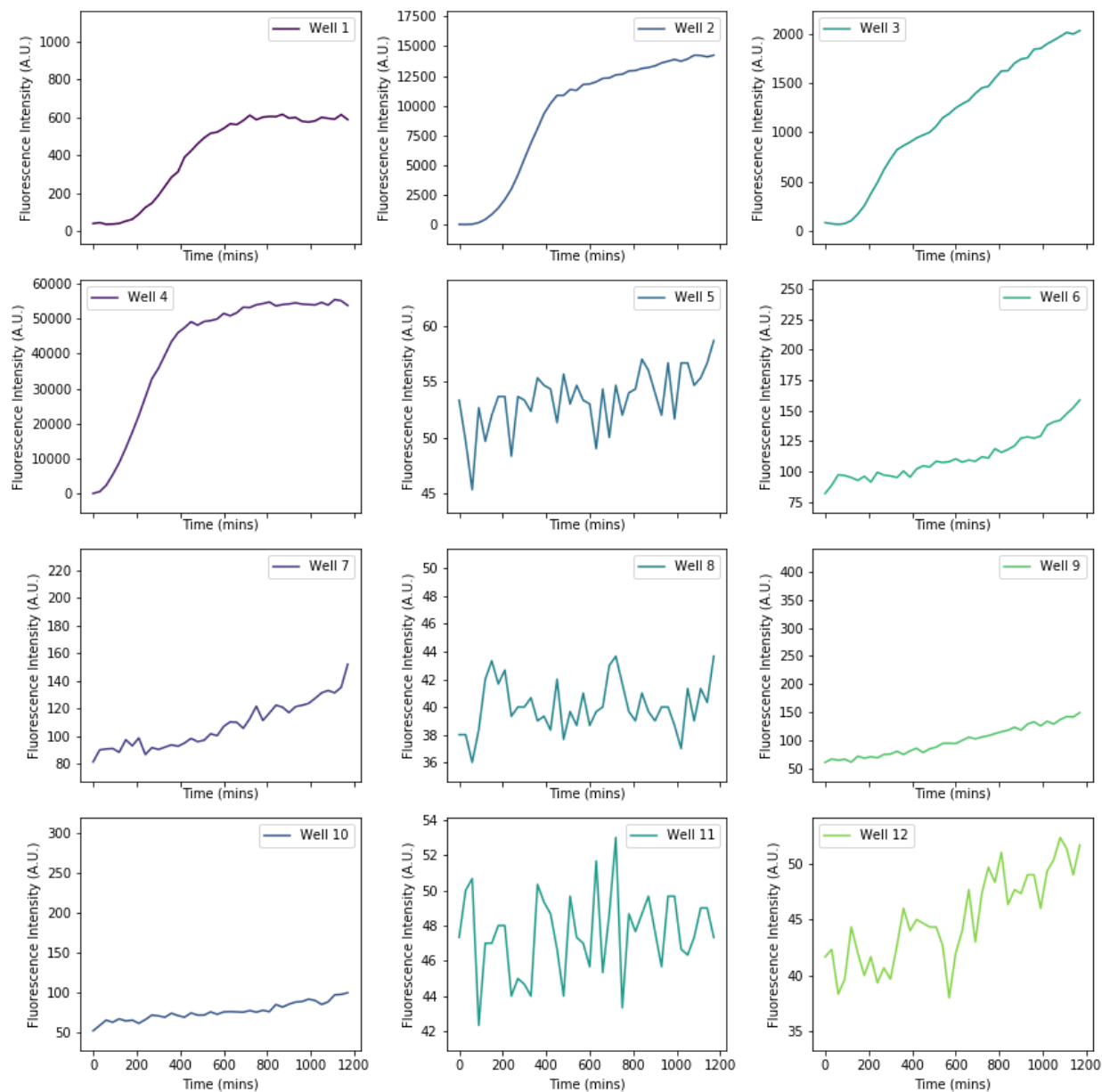


Figure 3.24 Kinetics of deGFP-BMP2 over a 20-hour time-lapse, from an area furthest away from where the bioink gels were placed in a 96-well plate (Position 2). Average of 3 technical repeats. Well 1 refers to a 1:1 diluted negative CFES (-DNA) with DEPC water. Well 2: a 1:1 diluted positive CFES (+DNA) with DEPC water. Well 3: a 1:3 diluted negative CFES (-DNA) with DEPC water. Well 4: a 1:3 diluted positive CFES (+DNA) with DEPC water. Well 5: a 1:1 v/v of bioink (final 6% w/v alginate and 13% w/v F127) with DEPC water. Well 6: a 1:1 v/v of bioink with +ve CFES in 100 mM CaCl₂. Well 7: a 1:1 v/v of bioink with -ve CFES in 100 mM CaCl₂. Well 8: a 1:3 v/v of bioink (final 6% w/v alginate and 13% w/v F127) with DEPC water. Well 9: a 1:3 v/v of bioink with +ve CFES. Well 10: a 1:3 v/v of bioink with -ve CFES. Wells 11 and 12: empty. 1 nM pdeGFP-BMP2 as input DNA. Excitation=479 ± 20 nm, emission= 520 ± 20 nm. Gain= 70. Temperature= 29°C. Extract 4.

3.4 CONCLUSIONS AND FURTHER WORK

In conclusion, cell-free protein synthesis systems, primarily in-house *E.coli* based extracts, were investigated and optimised to facilitate expression using a series of the BMP2 and VEGF modified plasmids. The work included pushing the boundaries of protein yield by probing calibration components per extract, reaction temperature, input DNA concentration or befitting ratios, addition of supplements such as T7 RNA polymerase, and hydrogels as chassis.

Throughout this chapter, pdeGFP served as an excellent benchmark for the other functional plasmids, due to its high fluorescence output. The series of pCellFree were one of the first successfully constructed circular DNA. It was shown that the products of some of the pCellFree plasmids were synthesised in PURExpress to the same level as deGFP with in-house CFES. By supplementing the in-house CFES with T7 RNA polymerase in a plasmid form (Pr1-T7RNAP), it was possible to significantly increase the yield of BMP2 modified pCellFree, even surpassing that of PURExpress. There are other plasmids encoding T7 RNAP with stronger promoters, such as Pr-T7RNAP with single base mutations, that could enhance the expression from pCellFree. Although site-directed, ligase-independent mutagenesis was attempted to change the promoter sequence, it was unsuccessful (not shown here). It would be compelling to pursue other T7 RNAP plasmids with stronger promoters. Thus far, utilising T7 RNAP in an enzyme form with the in-house system was ineffective; this could be due to the absence of the transcriptional buffer in those reactions. The commercial buffer contains reagents that are not present in the in-house CFES, but could be beneficial to the activity of the polymerase. To simplify the mode of transcription of pCellFree with in-house CFES, the addition of the transcriptional buffer with T7 RNAP could be explored. Moreover, finding a more befitting concentration of the T7 RNAP enzyme than explored here, could also yield a positive result.

A two-plasmid system of pBEST-p15A-OR2-OR1-Pr-UTR1-Sigma28-T500 and pTar-BMP2-EGFP was also investigated for functionality with in-house CFES. The setup

proved to be effective when the ratios of sigma28 : pTar-BMP2-EGFP were arranged at 1:10, 1:25 and 1:35. Although the fluorescence levels were low in comparison to deGFP, they were higher than for the original unmodified AqpZ-EGFP. There also appears to be variation in expression levels between DNA originating from different bacterial colonies. The ratio optimization should be investigated further, as no dominant setting was established. The major advantage of implementing a two-plasmid system within CFES is the ability to create a transcriptional trigger which can be further incorporated into more complex synthetic gene networks.

The most high-yielding plasmid that was designed was the pdeGFP-BMP2, with 0.49 mg/mL protein produced in a 10 μ L reaction at 29°C. DNA threshold where the efficiency of transcription was the highest for pdeGFP-BMP2 was determined to be at 3 nM. It was also established that protein synthesis was significantly impacted by the temperature of the reaction, with 29°C being more favourable than 37°C. pdeGFP-BMP2 was chosen for further experimentation detailed in following chapters.

The preliminary experiments of in-house CFES with bioink components sodium alginate and Pluronic-F127 showed promising evidence of deGFP production, demonstrating the ability to create a protein-producing hydrogel. When the complete bioink formulation was investigated for CFES, some increase in fluorescence was observed, although, large error bars prevented from distinguishing significance. The output from agar and agarose CFES was higher in comparison, however, the problem of sample variance was persistent. This was a reoccurring theme when the fluorescence diffusion experiments were monitored. For CFES containing reactions, a lower signal in the surrounding liquid than in the gel was observed. However, there was a minimal difference between positive CFES gels and negative CFES gels. The limited experimental setup made it difficult to assess if the protein was produced from the gel itself or perhaps from CFES reaction that did not incorporate into the gel. Confocal microscopy could be performed to establish the homogeneity of the fluorescence across the gel. Additionally, a more thorough method of

mixing between CFES and gel components should be carried out to generate consistent readings. A reaction scale-up could help with that, as well as mitigate the evaporation losses noted during the experiment. Another way to aid consistent readings could be achieved by controlling the same temperature between the gels and CFES components upon mixing. This would prevent the gelation triggered by temperature difference before the components are integrated within the gel. Other possible bioinks such as hyaluronic acid⁴² or chitosan⁴³ could be considered for CFES. This is particularly important depending on the chosen route of protein expression and/or release trigger. Lyophilisation of in-house CFES could also be explored as a method to overcome the weakened fluorescence signal when CFES is diluted upon mixing with bioink components. Instead of using liquid components, freeze-dried CFES could be dissolved using the gels.

Overall, when CFES is considered, being able to monitor the separate dynamics of mRNA and protein synthesis would provide an improved understanding of the system. It may help identify bottlenecks at both steps of TX-TL, which are important for engineering, including gene circuits. There are numerous ways of tracking mRNA transcription, such as radioactive labelling⁴⁴, the use of molecular beacons⁴⁵, or fluorescent aptamers⁴⁶, implementing one of those would greatly add to the real-time monitoring toolkit. By identifying the rate-limiting steps, optimisation of the input DNA or the reaction components could be performed. This could take a form of supplementing the reaction with elongation factors, chaperones, or secondary energy sources, amongst others. This fine-tuning of the components could significantly elevate the efficiency of CFES. One other aspect of CFES that is worth exploring is the use of a different cell extract. The active form of a native BMP2 is a dimer, which consequently means that disulfide bonds need to be formed. However, the reducing environment of a bacterial cytoplasm prevents the formation of disulfide bonds and therefore, becomes problematic for efficient folding of recombinant proteins⁴⁷. This effect is reduced for CFES, where only the cell extract is used, but there are certain extract, such as the Origami *E.coli* strains which enhance

disulfide bond formation⁴⁸. The addition of iodoacetamide to the cell extract lowers the reducing activity and therefore, should also be investigated⁴⁹.

3.5 BIBLIOGRAPHY

1. Olliver, C. L. & Boyd, C. D. In Vitro Translation of Messenger RNA in a Rabbit Reticulocyte Lysate Cell-Free System. in *Nucleic Acids* 145–156 (Humana Press, 1984). doi:10.1385/0-89603-064-4:145.
2. Harbers, M. Wheat germ systems for cell-free protein expression. *FEBS Letters* 588, 2762–2773 (2014).
3. Dortay, H. & Mueller-Roeber, B. A highly efficient pipeline for protein expression in *Leishmania tarentolae* using infrared fluorescence protein as marker. *Microbial Cell Factories* 9, 1–10 (2010).
4. Smolskaya, S., Logashina, Y. A. & Andreev, Y. A. Molecular Sciences Escherichia coli Extract-Based Cell-Free Expression System as an Alternative for Difficult-to-Obtain Protein Biosynthesis. *International Journal of Molecular Sciences* 21, (2020).
5. Tuckey, C., Asahara, H., Zhou, Y. & Chong, S. Protein Synthesis Using A Reconstituted Cell-Free System. *Current protocols in molecular biology* 108, 1–22 (2014).
6. Sun, Z. Z. *et al.* Protocols for Implementing an Escherichia coli Based TX-TL Cell-Free Expression System for Synthetic Biology. *Journal of Visualized Experiments* 1–15 (2013) doi:10.3791/50762.
7. Zubay, G. In vitro synthesis of protein in microbial systems. *Annual Review of Genetics* 7, 267–287 (2003).
8. Jia, B. & Jeon, C. O. High-throughput recombinant protein expression in Escherichia coli: current status and future perspectives. *Open Biology* 6, (2016).

9. Shrestha, P., Holland, T. M. & Bundy, B. C. Streamlined extract preparation for Escherichia coli-based cell-free protein synthesis by sonication or bead vortex mixing. *Biotechniques* 53, 163–174 (2012).
10. Shin, J. & Noireaux, V. Efficient cell-free expression with the endogenous E. Coli RNA polymerase and sigma factor 70. *Journal of Biological Engineering* 4, (2010).
11. Marshall, R. & Noireaux, V. Quantitative modeling of transcription and translation of an all-E. coli cell-free system. *Scientific Reports* 2019 9:1 9, 1–12 (2019).
12. Ferrara, N., Gerber, H.-P. & LeCouter, J. The biology of VEGF and its receptors. *Nature Medicine* 9, 669–676 (2003).
13. Chen, D., Zhao, M. & Mundy, G. R. Bone Morphogenetic Proteins. *Growth Factors* 22, 233–241 (2004).
14. Phillips, A. M. Overview of the fracture healing cascade. *Injury* 36, S5–S7 (2005).
15. Gagoski, D. *et al.* Gateway-compatible vectors for high-throughput protein expression in pro- and eukaryotic cell-free systems. *Journal of Biotechnology* 195, 1–7 (2015).
16. Shin, J. & Noireaux, V. An E. coli cell-free expression toolbox: Application to synthetic gene circuits and artificial cells. *ACS Synthetic Biology* 1, 29–41 (2012).
17. Rogers, M. B., Shah, T. A. & Shaikh, N. N. Turning Bone Morphogenetic Protein 2 (BMP2) on and off in Mesenchymal Cells. *Journal of Cellular Biochemistry* 116, 2127–2138 (2015).
18. Caliarì, S. R. & Burdick, J. A. A practical guide to hydrogels for cell culture. *Nature Publishing Group* 13, 405 (2016).
19. Lee, J. H. Injectable hydrogels delivering therapeutic agents for disease treatment and tissue engineering. *Biomaterials Research* 22, 1–14 (2018).

20. Zhou, X., Wu, H., Cui, M., Lai, S. N. & Zheng, B. Long-lived protein expression in hydrogel particles: Towards artificial cells. *Chemical Science* 9, 4275–4279 (2018).
21. Pardee, K. *et al.* Paper-Based Synthetic Gene Networks. *Cell* 159, 940–954 (2014).
22. Park, N. *et al.* High-yield cell-free protein production from P-gel. *Nature Protocols* 4, 1759–1770 (2009).
23. Whitfield, C. J. *et al.* Cell-free protein synthesis in hydrogel materials. *Chemical Communications* 56, 7108–7111 (2020).
24. Armstrong, J. P. K., Burke, M., Carter, B. M., Davis, S. A. & Perriman, A. W. 3D Bioprinting Using a Templated Porous Bioink. *Advanced Healthcare Materials* 5, 1724–1730 (2016).
25. Shimizu, Y. *et al.* Cell-free translation reconstituted with purified components. *Nature* 19, 751–755 (2001).
26. Leipply, D., Lambert, D. & Draper, D. E. Ion-RNA Interactions: Thermodynamic Analysis of the Effects of Mono- and Divalent Ions on RNA Conformational Equilibria. in *Methods in Enzymology* vol. 469 433–463 (2009).
27. Record, M. T. J., Courtenay Elizabeth S, Cayley Scott & Guttman Harry. Biophysical compensation mechanisms buffering E.coli protein-nucleic acid interactions against changing environments. *Cell Biology* 23, (1998).
28. Lambert, D., Leipply, D., Shiman, R. & Draper, D. E. The influence of monovalent cation size on the stability of RNA tertiary structures. *Journal of Molecular Biology* 390, 791–804 (2009).
29. Muller-Hill, B. *The lac Operon : a short history of a genetic paradigm.* (Walter de Gruyter, 1996).

30. Pelley, J. W. RNA Transcription and Control of Gene Expression. in *Elsevier's Integrated Review Biochemistry* 137–147 (W.B. Saunders, 2012). doi:10.1016/B978-0-323-07446-9.00016-7.
31. Ramos, J. L., García-Salamanca, A., Molina-Santiago, C. & Udaondo, Z. Operon. in *Brenner's Encyclopedia of Genetics: Second Edition* vol. 5 176–180 (Elsevier Inc., 2013).
32. Hansen, L. H., Knudsen, S. & Sørensen, S. J. The Effect of the lacY Gene on the Induction of IPTG Inducible Promoters, Studied in *Escherichia coli* and *Pseudomonas fluorescens*. *Current Microbiology* 36, 341–347 (1998).
33. Dickson, K. A., Haigis, M. C. & Raines, R. T. Ribonuclease Inhibitor: Structure and Function. *Prog Nucleic Acid Res Mol Biol* 80, 349–374 (2005).
34. Siegal-Gaskins, D., Tuza, Z. A., Kim, J., Noireaux, V. & Murray, R. M. Gene circuit performance characterization and resource usage in a cell-free “breadboard.” *ACS Synthetic Biology* 3, 416–425 (2014).
35. Coutable, A. *et al.* Preparation of tethered-lipid bilayers on gold surfaces for the incorporation of integral membrane proteins synthesized by cell-free expression. *Langmuir* 30, 3132–3141 (2014).
36. Calamita, G. The *Escherichia coli* aquaporin-Z water channel. *Molecular Microbiology* 37, 254–262 (2000).
37. Schneider, B. *et al.* Membrane Protein Expression in Cell-Free Systems. *Methods Mol Biol* 601, 165–186 (2010).
38. Garenne, D. *et al.* Cell-free gene expression. *Nature Reviews* 1, (2021).

39. Kim, J., Copeland, C. E., Seki, K., Vögeli, B. & Kwon, Y. C. Tuning the Cell-Free Protein Synthesis System for Biomanufacturing of Monomeric Human Filaggrin. *Frontiers in Bioengineering and Biotechnology* 8, 1244 (2020).
40. Gagoski, D. *et al.* Performance benchmarking of four cell-free protein expression systems. *Biotechnology and Bioengineering* 113, 292–300 (2016).
41. Ge, X., Luo, D. & Xu, J. Cell-Free Protein Expression under Macromolecular Crowding Conditions. *PLoS one* 6, (2011).
42. Petta, D., Ambrosio, L., Grijpma, D. W. & Eglin, D. Hyaluronic acid as a bioink for extrusion-based 3D printing. *Biofabrication* 12, (2020).
43. Ku, J. *et al.* Cell-Laden Thermosensitive Chitosan Hydrogel Bioinks for 3D Bioprinting Applications. *Applied Sciences* 10, (2020).
44. Porecha, R. & Herschlag, D. RNA radiolabeling. in *Methods in Enzymology* vol. 530 255–279 (Academic Press Inc., 2013).
45. Monroy-Contreras, R. & Vaca, L. Molecular Beacons: Powerful Tools for Imaging RNA in Living Cells. *Research Journal of Nucleic Acids* 2011, 1–15 (2011).
46. Dolgosheina, E. v *et al.* RNA Mango Aptamer-Fluorophore: A Bright, High-Affinity Complex for RNA Labeling and Tracking. *ACS Chemical Biology* 9, 2412–2420 (2014).
47. Stewart, E. J. & Beckwith, J. Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins. *The EMBO Journal* 17, 5543–5550 (1998).
48. Xiong, S. *et al.* Solubility of disulfide-bonded proteins in the cytoplasm of Escherichia coli and its “oxidizing” mutant. *World Journal of Gastroenterology* 11, 1082 (2005).

Chapter 3 - Optimisation of the Cell-Free Expression System

49. Yin, G. & Swartz, J. R. Enhancing Multiple Disulfide Bonded Protein Folding in a Cell-Free System. *Biotechnology and Bioengineering* 86, 188–195 (2004).

CHAPTER 4. CHARACTERISATION OF THE DEGFP-BMP2 CHIMERA

4.1 INTRODUCTION

Following the protein yield comparisons between various plasmids used in the CFES (Chapter 3), the deGFP-BMP2 plasmid was determined to be the best candidate for further experimentation. Recombinant human BMP2 fused with a fluorescent protein is not commercially available, and this particular chimera of deGFP-BMP2 adapted for cell-free expression is a novel fusion protein. Studies have been conducted where BMP2 was labelled with dyes such as Alexa Fluor¹, Dylight². Moreover, BMP2 has been expressed in *E.coli* as an N-terminus extension of a full length wtGFP, for a retention assay³. These studies showed no reduction in BMP2 activity as an effect of the labelling, or significant changes in the fluorescent properties of GFP. It was necessary to conduct full characterisation of deGFP-BMP2 due to the use of a CFES-adapted plasmid backbone and a minimal domain fluorophore deGFP, all of which can play a part in the folding, dimerisation and subsequent function of the chimera. To determine the effect of fusing deGFP with BMP2, the chimera was expressed in *E.coli* and purified to obtain significant quantities, sufficient for the analysis of the secondary and the tertiary structure of the chimera. The data was compared with an exemplary fluorophore, enhanced green fluorescent protein (EGFP).

EGFP structure is similar to that of wild type Green Fluorescent Protein, with the emblematic β -barrel structure encasing the chromophore within and an α helix spanning the core⁴. It has been reported that the β -barrel, consisting of 11 strands, constitutes 47% of the secondary structure of EGFP, whereas 13% of the protein secondary structure is helical, composed of 3_{10} -helix and α -helix conformations. The remaining 40% comes from the loops at the ends of the barrel forming the coil configurations⁵. When it comes to amino acid deletions in EGFP, certain variants can imbue positive characteristics such as increased folding, dynamics and fluorescence driven by alterations in the molecular structure, whereas other amino acid deletions can be detrimental⁶. deGFP used in this thesis is a variant of EGFP with the minimal domain of amino acids 6-229 (Figure 4.1).

The minimal domain is the sequence that is crucial to maintain fluorescence of EGFP. The 15 amino acids that were deleted in this variant included the small α -helix at the N-terminus and a short tail just after the last β -sheet at the C-terminus⁷. Some of the crucial conformations of the minimal domain include the two large loops on the ends of the barrel located at amino acids 129-142 and 189-196, two small α -helices at amino acids 76-81 and 83-88, the central α -helix as well as the surrounding β -sheets completing the barrel⁸. Contrastingly, BMP2, which contributes as the other part to the chimera, forms more of a ‘butterfly’ or ‘two fists’ shape when it dimerises.

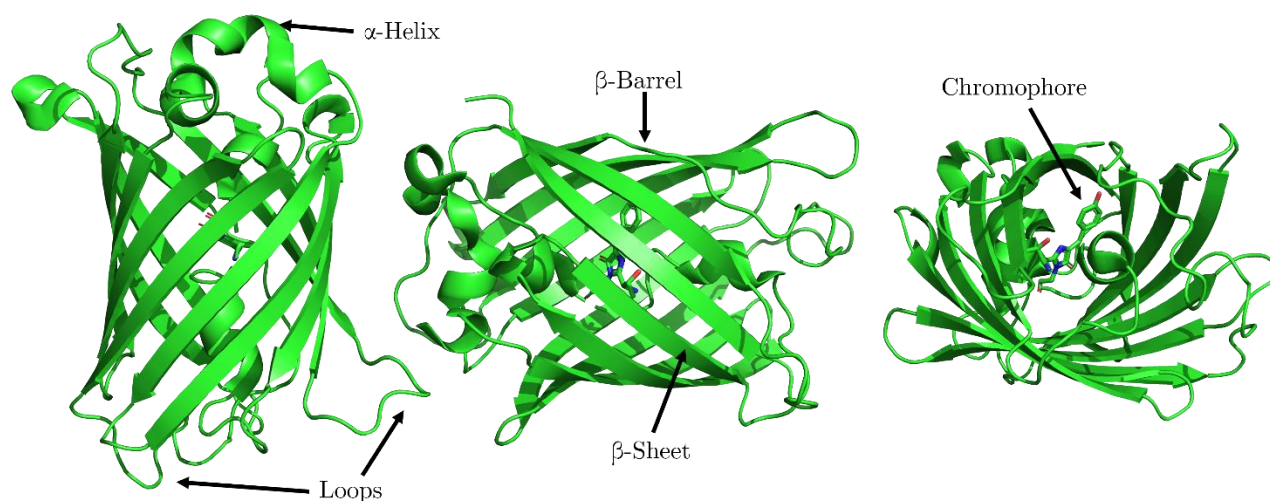


Figure 4.1 Schematic representation of the hypothesised deGFP at three various angles. Figure was generated from EGFP PDB ID: 2Y0G, by displaying amino acid residues 6-229 only.

In a mammalian cell, BMP2 is synthesised as a large precursor with a signal peptide at the amino terminal, a pro-domain, and a mature sequence at the carboxyl terminal. During the processing, the signal peptide is removed to create a pro-BMP2, which is then proteolytically cleaved by proteases known as proprotein convertases, whilst two mature BMP2 sequences dimerise *via* the cysteine knot, which is made possible by seven highly conserved cysteines⁹. Since the work here is carried out in CFES or a bacterial host, the signal peptide for transport processing *via* the endoplasmic reticulum and the trans Golgi network is unnecessary, consequently, only the mature BMP2 was included in the primary sequence of the chimera (Positions 283-396). The secondary structure of BMP2 has been

reported to consist of 10 β -sheets that define the two finger-like elements of the folding topology, a four-turn helix contributing to 14% of the structure and forming the ‘wrist’ of the hand (Figure 4.2 A), and completing the structure with the ‘palm of the hand’ formation is the stabilising cysteine knot positioned at Cys296-Cys361, Cys325 with Cys393 and Cys329-Cys395¹⁰ (Figure 4.2 C). During the formation of the homodimer, the monomers join at the conserved cysteine residue at position 360 creating a disulphide bridge, with the β -strands bending to generate a concave and a convex surface for the two types of receptor interactions¹¹ (Figure 4.2 B)

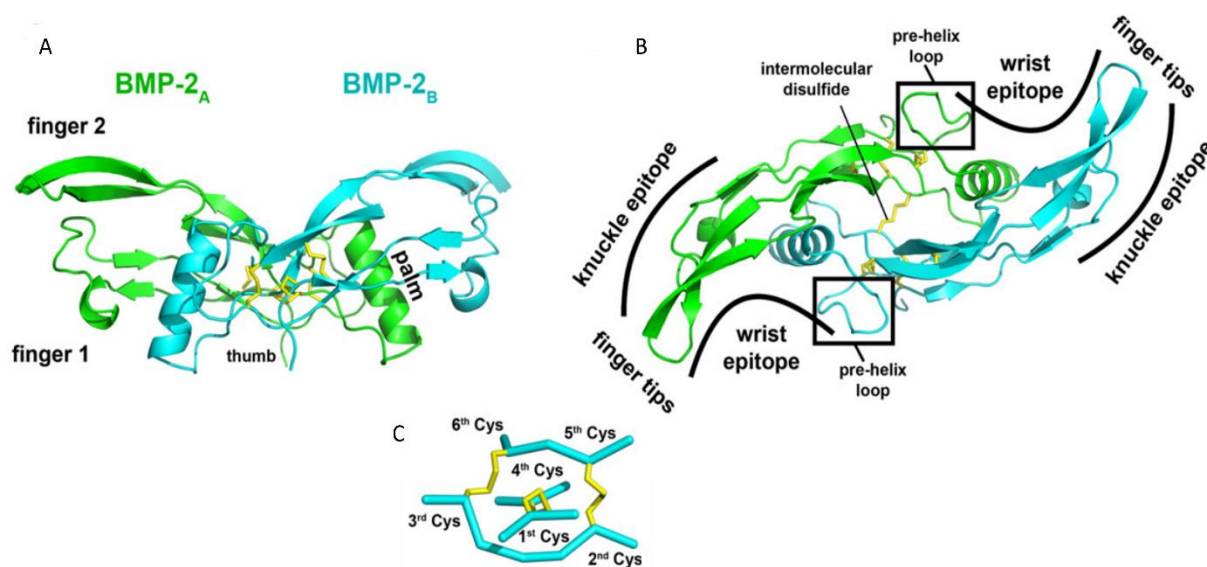


Figure 4.2 Ribbon representation of BMP2 dimer formed from two monomers labelled in green and blue. A: The ‘butterfly’ or ‘fists’ representation of the dimer with two finger-like topology generated by the β -sheets, N-terminus represented by the thumb, the helix forming the wrist and the monomer to monomer interface generates the palm. B: View of the BMP2 dimer along the two-fold symmetry axis. C: Cysteine knot formed by the disulphide bridges. Reproduced from Mueller *et al.*¹² Here, the arrangement of the 352 amino acids of the chimera begins at deGFP, followed by a TEV cleavage site, mature sequence of BMP2 and a 6xHis-tag at the C-terminus of the BMP2 (Figure 4.3 A). It is hypothesised that upon dimerisation, the monomers of BMP2 join together forming a disulphide bridge, and each monomer carries a singular deGFP molecule attached to its N-terminus (Figure 4.3 B).

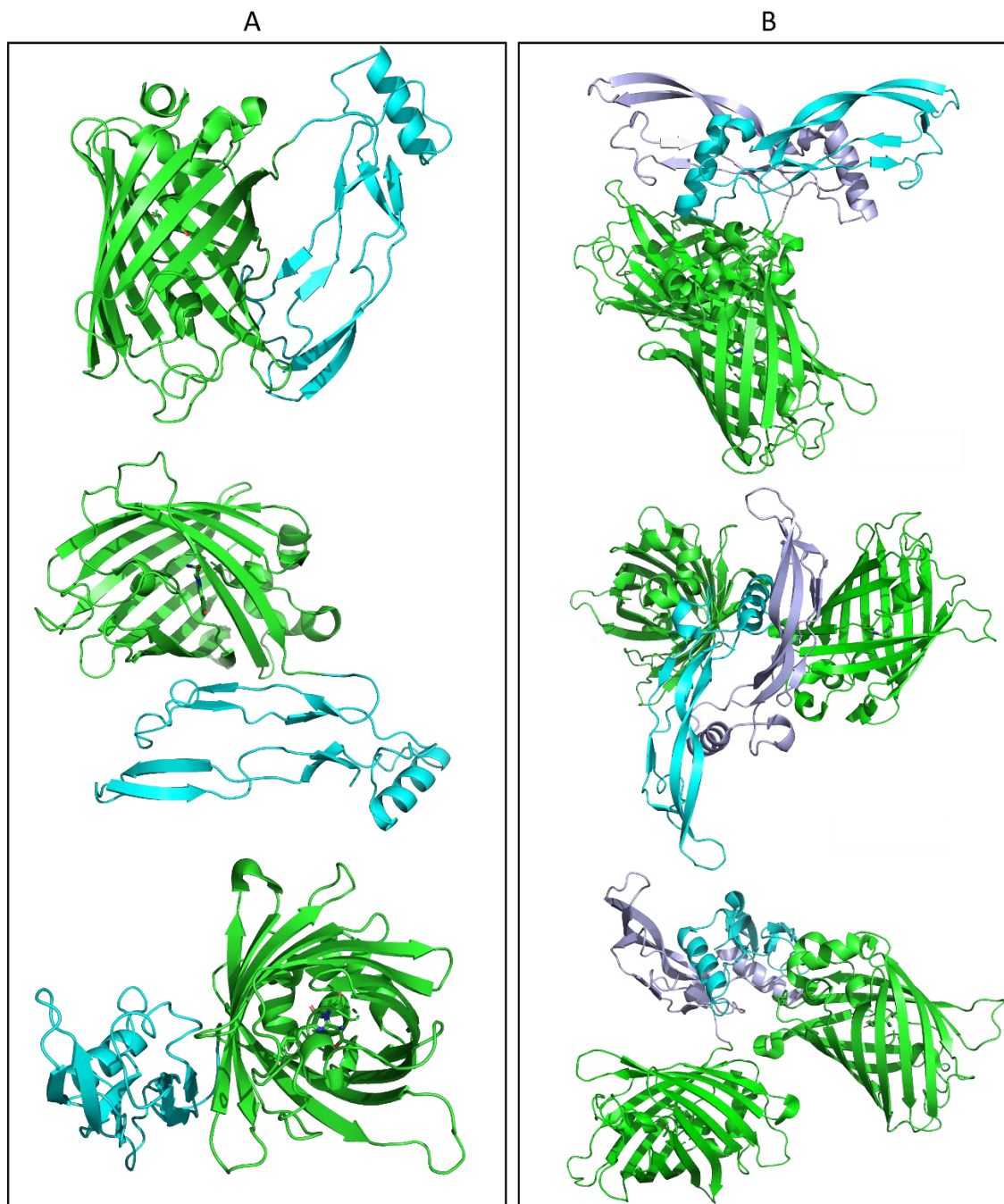


Figure 4.3 Schematic representation of the hypothesised deGFP-BMP2 chimera at three various angles. A: Monomeric chimera. N-terminus of BMP2 monomer (cyan) PDB ID: 3BMP joined to the C-terminus of modified EGFP (green) PDB ID: 2Y0G (showing only the minimal domain of 6-229 amino acids). B: Dimeric chimera. BMP2 dimer (cyan and lilac) PDB ID: 1ES7 with ectodomains of receptors not shown, with each monomer joined to the C-terminus of modified EGFP (green) PDB ID: 2Y0G (showing only the minimal domain of 6-229 amino acids). TEV cleavage site between BMP2 and deGFP, and the 6xHis-tag at the C-terminus of BMP2 are omitted in both versions.

Chapter 4 - Characterisation of the deGFP-BMP2 Chimera

In this chapter, *E.coli* expression and purification of deGFP-BMP2 is described, followed by the investigation of the secondary and tertiary structure of this chimera, using techniques including circular dichroism, UV-Vis and fluorescence spectroscopy as well as dynamic light scattering. The resulting analyses show secondary structure of the chimera that is consistent with the constituent proteins and the data points to the presence of the hypothesised dimeric BMP2 with two deGFP molecules.

4.2 MATERIALS AND METHODS

4.2.1 SMALL SCALE EXPRESSION OF DEGFP-BMP2 IN *E. COLI*

A 2 μ L aliquot of deGFP-BMP2 plasmid was added to 50 μ L *E.coli* competent cells of either Rosetta 2 (Novagen)/ SHuffle T7 (New England Biolabs)/ BL21 (New England Biolabs) strain and incubated on ice for 30 minutes. The DNA-cells mixture was heat shocked at 42°C for 45 seconds and then immediately transferred to ice for 2 minutes. Of the pre-warmed SOC media (New England Biolabs), 950 μ L was added to the cells and placed in a shaking incubator at 37°C, 200 rpm for 1 hour before plating onto yeast extract with tryptone (2xYT) for Rosetta 2, or Luria broth (LB) for Shuffle T7 and BL21 agar plates containing carbenicillin for overnight incubation at 37°C. Single colonies were selected for overnight growth in 5 mL 2xYT+P/LB media containing carbenicillin at 37°C with rotation at 200 rpm. For the small-scale expression, 1 mL of the overnight starter culture was suspended in 20 mL of fresh autoclaved media and incubated again at 37°C with rotation at 200 rpm until the optical density measured at 600 nm reached 0.6-0.8. In the induction experiment, 1 mM as the final concentration of IPTG was added to some cultures and incubated again for 3 hours. This approach was utilised to establish the constitutive expression of the plasmid, rather than an induction mechanism being at play for its expression. For the temperature and time experiment, the 20 mL bacterial cultures were incubated at either 25°C or 37°C for 24 hours or 48 hours. Cells were collected by centrifugation at 4000 g, 4°C for 15 minutes using Megafuge ST Plus (Thermo Fisher Scientific). The cell pellet was resuspended with 5 μ L per mg of wet cell pellet BugBuster Protein Extraction Reagent (Novagen) by pipetting. Phylmethylsulfonyl fluoride (PMSF) was added at 1 μ L per mL of BugBuster used, and the resuspension was placed on a shaking platform for 15 minutes at room temperature. The insoluble cell debris was separated by centrifugation at 16,000 g for 20 minutes at 4°C. Separate fractions were tested with polyacrylamide gel electrophoresis (SDS-PAGE).

4.2.2 LARGE SCALE EXPRESSION OF DEGFP-BMP2 IN *E. COLI*

The transformation of Rosetta 2 *E. coli* cells with deGFP-BMP2 plasmid was performed in the same manner as stated in Section 4.2.1. Some of the grown colonies were selected and placed in 10 mL of yeast extract rich and tryptone media with phosphates, 2xYT+P containing carbenicillin, for overnight growth at 37°C, 200 rpm. The following day, 2 mL of the bacterial starter culture was used to inoculate each 1 L of fresh, autoclaved 2xYT+P media with carbenicillin, and placed in the shaking incubator at 25°C for 48 hours. IPTG was omitted in the large scale expression as the small scale experiment showed no increase in protein yield upon its addition, confirming there is no induction mechanism. Cells were harvested by centrifugation at 5000 rpm, 20 minutes at 4°C using Sorvall RC5C centrifuge with SLA-3000 rotor (DuPont). The media was removed, and the cell pellets were transferred into 50 mL falcon tubes and resuspended in 20 mM Tris HCl, 500 mM NaCl, 0.1% Triton X-100, pH 8.0. Cell lysis was performed by sonication with 1 second ‘on’ and 2 seconds ‘off’ alternation at 65% amplitude for a total of 6 minutes, with the falcon tube placed on ice. The lysed solution was centrifuged at 15,000 rpm, 4°C for 15 minutes, using Avanti J26XP centrifuge with JA-25.50 rotor (Beckman Coulter). To isolate the inclusion bodies, the pellet was subject to resuspension in 20 mM Tris HCl, 500 mM NaCl, 0.1% Triton X-100, pH 8.0, sonication, and centrifugation three more times. The final pellet was incubated with a stir bar in a buffer containing 20 mM Tris HCl, 150 mM NaCl, 20 mM imidazole, 5 mM dithiothreitol (DTT) and 6 M urea, pH 7.5, at 4°C overnight. The dissolved pellet solution was centrifuged at 15,000 rpm, 4°C for 45 minutes. The supernatant with the now soluble unfolded chimera was filtered through 0.45 µm and then 0.22 µm cellulose membrane syringe filters before purification using immobilised metal affinity chromatography (IMAC) (Section 4.2.3.1). To stimulate BMP2 refolding, the fractions of the purified protein were pooled into 10 mL, and added dropwise, at 4°C, into refolding buffer (55 mM Tris, 10.56 mM NaCl, 0.44 mM KCl, 550 mM Guanidine HCl, 2.2 mM MgCl₂, 2.2 mM CaCl₂, 550 mM L-arginine, 1 mM DTT, pH 8.2). The solution

was stirred at 4°C overnight. The following day, the solution was buffer exchanged, at 4°C, into a working buffer (20 mM Tris HCl, 150 mM NaCl, 20 mM imidazole, 4 M urea, pH 8.0) using 3.5K molecular weight cut off (MWCO) regenerated cellulose dialysis tubing (Thermo Fisher Scientific). The resulting protein was subjected to IMAC again. The resulting fractions were pooled and concentrated into 10 mL using 10K MWCO spin concentrator (Sartorius), before attempting further purification of monomers and dimers by Size Exclusion Chromatography (SEC) (Section 4.2.3.2).

4.2.3 PURIFICATION OF DEGFP-BMP2

4.2.3.1 IMMOBILISED METAL AFFINITY CHROMATOGRAPHY (IMAC)

This selective purification technique, first adapted by Everson and Parker¹³ and later popularised by Porath *et al.* in 1975¹⁴, relies on the interactions between the metal ions that are immobilised on a chromatographic support such as Sepharose, and the lone pair of electrons present on nitrogen or oxygen in amino acid residues like histidine¹⁵. The most commonly used metal ions include transition-metal ions such as Cu²⁺, Zn²⁺, Co²⁺, Fe²⁺ and Ni²⁺ which are considered as electron pair acceptors and participate in forming metal chelates. Hochuli *et al.* developed a quadridentate chelating agent, nitrilotriacetic acid (NTA) that provides a stronger and more selective binding with oligo-histidine, allowing for efficient purification of recombinant proteins engineered with a His-tag¹⁶. The removal of the recombinant protein from the Ni-NTA complex involves the elution with, for example, imidazole buffer of a stronger ionic strength which competes for the ligand exchange.

Here, HisTrap FF crude 5 mL pre-packed column (GE Healthcare) along with ÄKTA start chromatography system (GE Healthcare) was used to purify the solubilised deGFP-BMP2 chimera that contained an engineered His-tag. The column was washed with 5 column volumes of degassed MiliQ water, followed by 5 column volumes of a binding

buffer, in this case, 20 mM Tris HCl, 150 mM NaCl, 20 mM imidazole, 5 mM DTT and 6 M urea, pH 7.5. Once the column was equilibrated, the filtered sample was applied, followed by the binding buffer wash to remove unbound protein. Finally, the recombinant protein was eluted with a high imidazole buffer (20 mM Tris HCl, 150 mM NaCl, 300 mM imidazole, 5 mM DTT and 6 M urea, pH 7.5). The fractions were collected in 2 mL Eppendorf tubes and analysed by SDS PAGE. IMAC purification was used again after deGFP-BMP2 underwent refolding and dialysis. These two steps generated a large volume of the sample, which had to be concentrated for size exclusion chromatography. The same IMAC settings were applied, but this time, the binding buffer contained 20 mM Tris HCl, 150 mM NaCl, 20 mM imidazole, and 4 M urea, pH 8.0, whereas the elution was performed with buffer containing 20 mM Tris HCl, 150 mM NaCl, 300 mM imidazole, and 4 M urea, pH 8.0.

4.2.3.2 SIZE EXCLUSION CHROMATOGRAPHY (SEC)

Size exclusion chromatography is a separation technique of molecules in solution based on size or hydrodynamic volume, with the use of porous packing¹⁷. It was first introduced by Lathe and Ruthven in 1955, where starch was used as the matrix.¹⁸ Common modern gel-filtration matrices include agarose, dextran and polyacrylamide, each with differently sized pores in the beads, and therefore, different separation resolutions¹⁹. When the sample is introduced into the column, small molecules will have a greater permeation into the pores than large molecules, therefore, due to the longer retention times, small molecules will elute later than larger molecules.

The SEC column, HiLoad 26/60 Superdex 200 pg (GE Healthcare) was firstly equilibrated with the sample buffer (20 mM Tris HCl, 150 mM NaCl, 4 M urea, pH 8.0). Before the sample was injected, it was concentrated to 10 mL using 10K MWCO spin concentrator (Sartorius). After sample injection, the column was subject to more sample buffer to aid elution. Fractions of 2 mL were generated at a flow rate of 1 mL/min. Bio-Rad NGC™ system was used as means of automating valve switching and monitoring.

4.2.4 POLYACRYLAMIDE GEL ELECTROPHORESIS (PAGE)

Polyacrylamide gel electrophoresis (PAGE) is a method of using an electrical field to separate molecules based on their molecular weight²⁰. Typically, the method involves an input of samples into a gel matrix, such as polyacrylamide submerged in a buffer and placed in between two electrodes. The separation of molecules in the gel matrix is influenced by factors such as protein 3D structure and charge. In sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE), the non-covalent interactions are removed by the anionic SDS, whereas a reducing agent, β -mercaptoethanol, disrupts any disulphide bonds present in the protein, allowing the migration through the matrix to be based on polypeptide length²¹. When an electric field is applied, the smaller peptides can travel further towards the anode electrode at the bottom, whereas larger molecules are more likely to stay at the cathode electrode at the top. It is also possible to examine protein separation without disulphide bonds being broken under non-reducing conditions (Non-reducing/LDS PAGE), or in a protein's non-denaturing state (Native PAGE) depending on the buffer's contents²².

For SDS PAGE, 10 μ L of sample was mixed with 10 μ L of SDS sample buffer (400 mM glycerol, 50 mM EDTA, 0.5 mM Tris, 5% SDS, 7 mM β -mercaptoethanol), and heated to 95°C for 5 minutes. 10 μ L of the mixture was loaded into individual wells of a Novex Wedgewell Tris-Glycine gel at either 8-16% or 16% (Thermo Fisher Scientific) which was positioned in a Mini Gel Tank (Thermo Fisher Scientific) and submerged in 1x SDS Tris-Glycine running buffer (Thermo Fisher Scientific). For non-reducing conditions, 3 μ L of LDS sample buffer (Thermo Fisher Scientific) was mixed with the samples instead. For Native PAGE, 5 μ L Native Novex sample buffer (Thermo Fisher Scientific) was mixed with the protein samples, which did not undergo the heating treatment, and the running buffer that was used was the Novex Tris-Glycine Native Running Buffer (Thermo Fisher Scientific). For any of the above conditions, 225 V and 125 mA was set on the power source for 35 minutes. The gel was removed from its casing and rinsed in MiliQ water

before staining for 30 minutes in Coomassie Brilliant Blue R-250 (Thermo Fisher Scientific). The dye was then removed, and the gel was de-stained in MiliQ water overnight.

4.2.5 CIRCULAR DICHROISM

Circular dichroism (CD) spectroscopy is a technique that facilitates determination of the chirality of molecules and analysis of the secondary structure and folding of proteins²³. Electromagnetic waves consist of an electric component and a magnetic field component that oscillate perpendicularly to the propagation direction of the light²⁴. In linearly polarised light, the electromagnetic wave oscillates along a single plane, whereas in circularly polarised light, the two components of the electromagnetic wave are at a 90° phase difference in relation to one another and therefore, the electric vector forms a helix along the direction of light propagation²⁵. For right circularly polarised light, the vector rotates clockwise in the frame of reference of the propagation direction, whereas for left circularly polarised light, the vector rotates counterclockwise. Optically active substances, also known as chiral, can give rise to CD signals, because of the difference in their absorbance of the circularly polarised light.

Here, samples were buffer exchanged with degassed CD buffer (10 mM K₃PO₄, 50 mM Na₂SO₄, pH 7.9) using 10K MWCO spin concentrator (Sartorius). The protein samples in correct buffer were filtered through a 0.22 µm syringe filter before being transferred to a CD sample cuvette at a concentration of 1.33 mg/mL. The data was collected using J-1500 CD spectrometer (Jasco), with the same settings for both EGFP and deGFP-BMP2 as stated in Table 4.1. The raw data was deconvoluted using BeStSel software²⁶.

Measurement range	260-190 nm
Temperature	25°C
Data pitch	1 nm
Direct integration time	4 s

Bandwidth	2 nm
Scanning speed	5 nm/min
Accumulations	3

Table 4.1. CD measurement settings.

4.2.6 ULTRAVIOLET-VISIBLE SPECTROSCOPY

Ultraviolet-visible (UV-vis) spectroscopy is an absorption spectroscopy technique most regularly used for quantifying solute concentrations due to it being directly proportional to the amount of absorbed light²⁷. The principle of this type of spectroscopy is based on the transition in electronic energy levels of molecules when a photon of light interacts with them when its amount of energy is the same as the energy difference between the levels. The absorption that corresponds to the electron promotion from the ground state to the excited state is measured as a function of wavelength by the spectrophotometer to produce a UV-vis spectrum²⁸. Accordingly, each molecule generates a unique fingerprint in UV-Vis due to its chemical makeup.

Protein samples of 0.35 mg/mL were subject to absorbance measurements using Cary 60 UV-Vis spectrometer (Agilent Technologies). The measurements were set according to Table 4.2 and performed in a 1 cm pathlength quartz cuvette. Beer-Lambert law was utilized to derive the concentrations of the proteins (Equation 2).

Measurement range	200-700 nm
Temperature	25°C
Scan rate	300 nm/min
Data interval	0.5 nm
Average time	0.1 s
Accumulations	3

Table 4.2. UV-vis measurement settings.

$$A = \text{Log}_{10} (I_0/I) = \epsilon \times C \times L$$

Equation 2. Beer-Lambert Law equation. A: absorbance, I_0 : intensity of the incident light at a given wavelength, I: transmitted intensity, ϵ : extinction coefficient, C: concentration, L: pathlength.

4.2.7 FLUORESCENCE SPECTROSCOPY

Fluorescence spectroscopy or simply, fluorimetry, is an electromagnetic spectroscopy technique of measuring fluorescence from certain compounds²⁹. When a photon is absorbed by the fluorophore, the energy of the molecule is raised to a higher, excited state but this is followed by relaxation where some vibrational energy is lost. Then, when the electrons in the lowest vibrational level of the excited electronic state drop back down to the ground electronic state, a photon of a lower energy is emitted³⁰. Therefore, an excitation spectrum is produced when the fluorophore absorbs photons at shorter wavelengths, and an emission spectrum is recorded at higher wavelengths when the lower energy photons are emitted.

Protein samples (0.01 mg/mL EGFP and 0.1 mg/mL deGFP-BMP2) in 20 mM Tris HCl, pH 8.0 were subject to excitation and emission spectra measurement taken with Cary Eclipse fluorometer (Agilent Technologies) according to Table 4.3.

Measurement range	350-700 nm
Temperature	25°C
Scan rate	120 nm/min
Data interval	1 nm
Average time	0.5 s
Accumulations	3
Excitation	488 nm
Emission	520 nm

Table 4.3. Fluorometry measurement settings.

4.2.8 DYNAMIC LIGHT SCATTERING

Dynamic Light Scattering (DLS) allows for determination of the particle size based on scattered light variations³¹. When a laser light irradiates a sample, the particles in suspension scatter the light in all directions, with the interference pattern between them dynamically changing due to their respective Brownian motion. Larger particles diffuse slower than smaller particles, leading to a higher degree of correlation between the signal at two closely spaced timepoints. The scattered light is collected by an orthogonal and a backscatter detector and the signal is converted using instrument in-built equations³².

EGFP in 25 mM HEPES, pH 7.0 and deGFP-BMP2 in 20 mM Tris-HCl, pH 8.0, were filtered through a 22 μm syringe filter before placing the samples in disposable cuvettes for measurements in ZetaSizer Nano ZS (Malvern Instruments). The parameters included 120 seconds equilibration time with three runs of each sample.

4.3 RESULTS AND DISCUSSION

4.3.1 EXPRESSION OF DEGFP-BMP2

To determine optimal *E.coli* recombinant expression conditions of deGFP-BMP2, a small scale culture with varying growing settings was carried out. Firstly, three different *E.coli* strains were transformed with the plasmid, including Rosetta 2, SHuffle and BL21. Rosetta 2 competent cells are BL21 derivatives that supply tRNAs with codons rarely used in *E.coli* and it is also the strain which was used to generate the in-house CFES extract³³. SHuffle stain was specifically engineered to allow for enhanced disulphide bond formation, which could prove useful in folding the multi-disulphide bonded BMP2³⁴. BL21 is the most common strain for recombinant protein expression, which lacks certain proteases for more stable product formation³⁵. In the small-scale expression, half of the cultures were induced with IPTG and the rest remained uninduced as ‘leaky’ (Figure 4.4). The induction with IPTG should not influence the expression yields since deGFP-BMP2 plasmid is not under a repressed promoter.

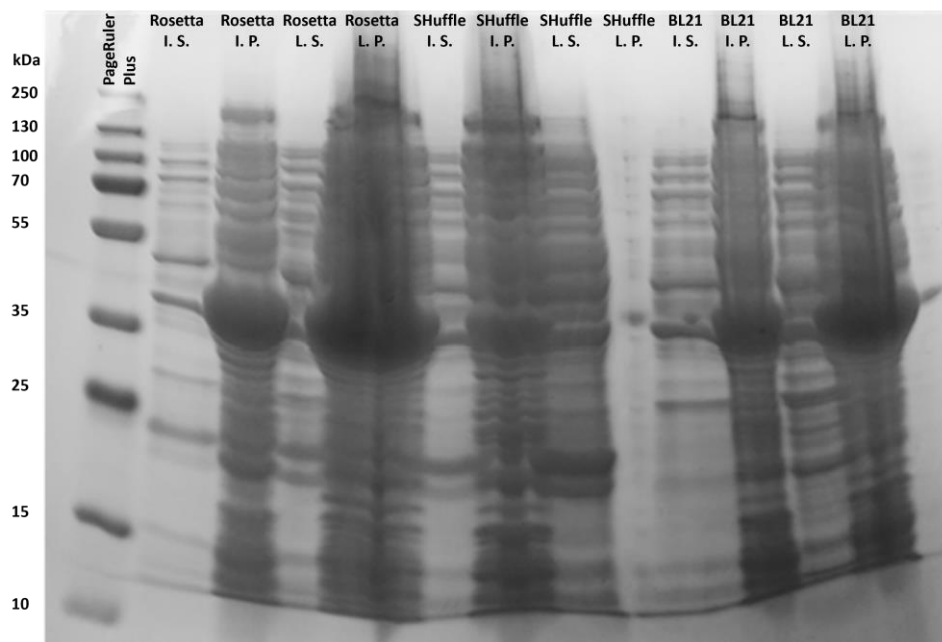


Figure 4.4 Reduced SDS PAGE of small-scale deGFP-BMP2 expression in Rosetta, SHuffle and BL21 bacterial cells. Expression was either induced with IPTG (I) or left uninduced (L). S= supernatant, P= pellet. PageRuler Plus as the protein ladder.

Chapter 4 - Characterisation of the deGFP-BMP2 Chimera

Following the expression of chimera in three different host strains, the cells were subject to protein extraction. Both the insoluble fraction and the supernatant were tested on SDS PAGE as seen in Figure 4.4. Although the lanes contained a large mixture of proteins in the non-purified lysate, an overexpression at around 40 kDa corresponding to the monomeric fusion could be observed. deGFP-BMP2 protein was mostly retained in the pellet which means further processes must be conducted to recover the protein from inclusion bodies. The highest protein yield has been noted in the Rosetta strain when 'leaky' expression was carried out (Rosetta L.P. in Figure 4.4). This strain has been chosen for further investigation into optimal culturing temperature and timing (Figure 4.5).

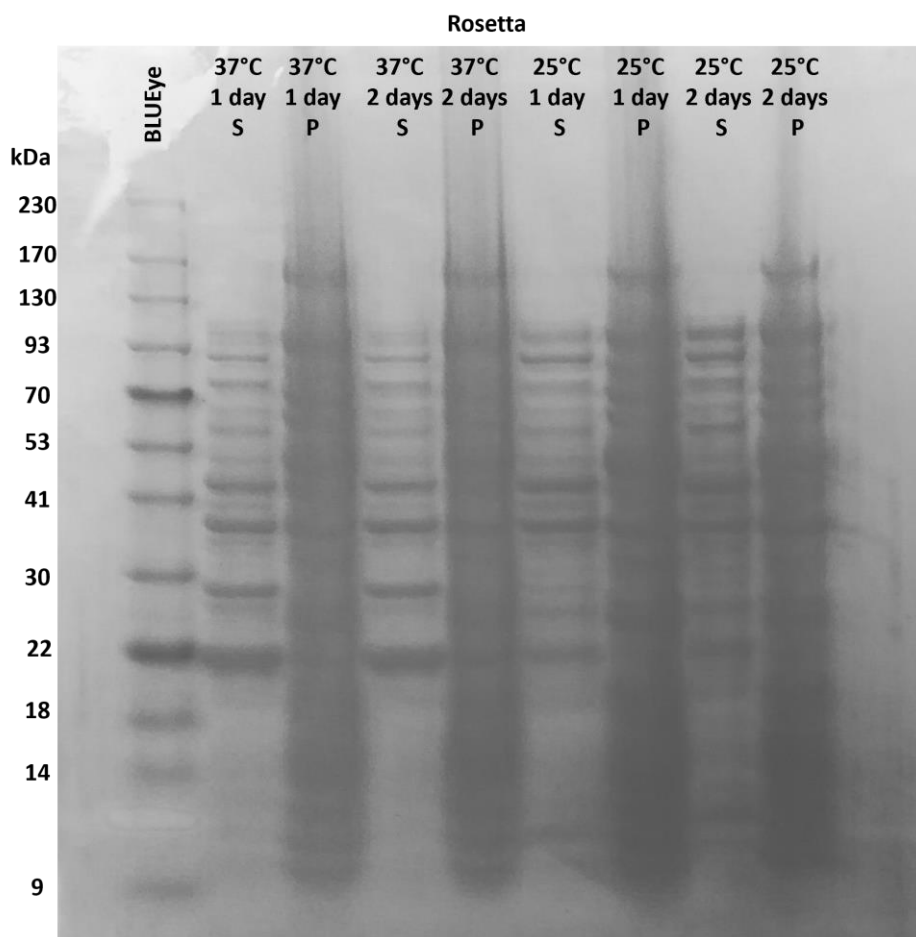


Figure 4.5 Reduced SDS PAGE of small-scale expression of deGFP-BMP2 in Rosetta cells taken place over 24 hours or 48 hours at 25°C or 37°C. Both cell supernatant (S) and cell pellet (P) were analysed. BLUEye as protein ladder.

It was noted that the most intense band of interest in Figure 4.4 was produced from a two-day culture at 25°C, which was utilised for the large-scale expression. During the inclusion bodies isolation, it was necessary to purify the deGFP-BMP2 protein before proceeding with further steps. HisTrap IMAC purification was chosen at first, due to the presence of a HisTag on the protein of interest. The chromatograph of the purification was obtained from ÄKTA start (Figure 4.6 A) and the collected fractions were analysed on non-reduced LDS PAGE (Figure 4.6 B). In the chromatograph, aside from the fractionated segment, another peak in the absorbance past 20 mL can be seen. This corresponds to the binding buffer wash step in the purification process, which removes any weakly bound proteins from the column, labeled here as B.B. In Figure 4.6 B, the protein of interest in a monomeric form was observed in large quantities throughout the tested fractions, as well as being present in the protein loading onto the column (C.L) and the binding buffer wash (B.B) steps. A small quantity of the dimeric fusion can be seen in some lanes at round 70 kDa. The F5 to F15 fractions were pooled together to aid refolding of the fusion protein before further purification using size exclusion chromatography (SEC) (Figure 4.7). These pooled fractions also appeared green, suggesting the presence of the fluorophore. Although under physiological conditions BMP2 exists as a dimer, some evidence suggests that bone formation can be induced by monomeric BMP2 alone³⁶, whereas others argue that only the dimeric form allows for the bioactivity of BMP2³⁷. Here, by performing SEC purification, further separation of the monomeric and dimeric forms was performed. From the chromatograph in Figure 4.7 A, two distinct increases in absorbance were noted: a broad peak past 120 mL and a much sharper peak at 310 mL. The first peak appears to be eluted after the void volume, suggesting the presence of aggregates is unlikely. The fractions corresponding to these peaks were pooled and concentrated. Both the concentrated sample and the filtrate (F) from the bottom of the spin concentrator were analysed on non-reduced LDS PAGE along

Chapter 4 - Characterisation of the deGFP-BMP2 Chimera

with the concentrated 'column loading' (C.L) from the previous IMAC purification (Figure 4.7 B).

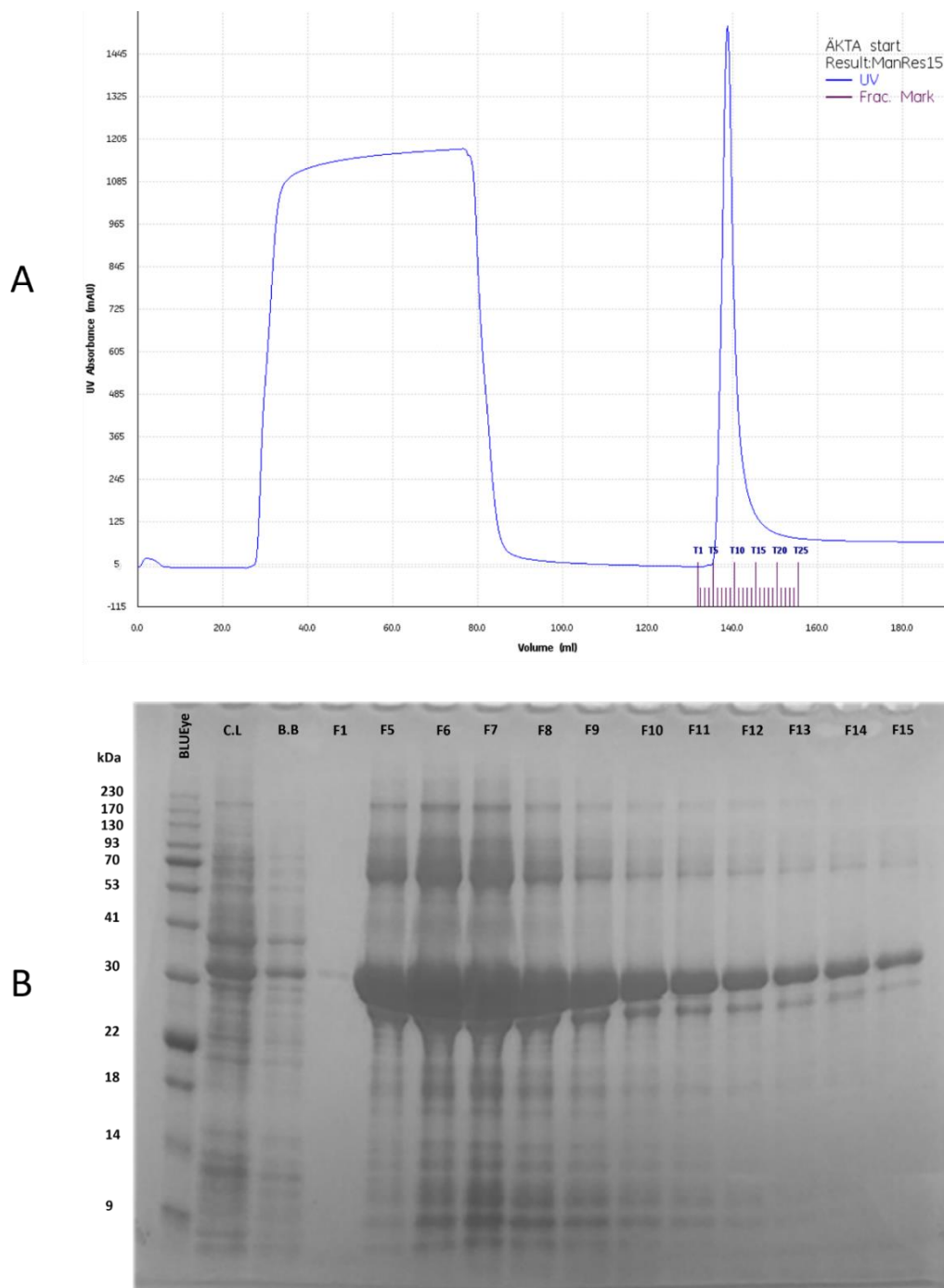


Figure 4.6 IMAC purification of deGFP-BMP2. A: chromatograph obtained from ÄKTA start system connected to HisTrap FF crude column. B: Non-reduced LDS PAGE of fractions eluted from HisTrap FF crude column. C.L: Column loading with the sample, B.B= Binding buffer wash off a protein loaded column. F: fraction.

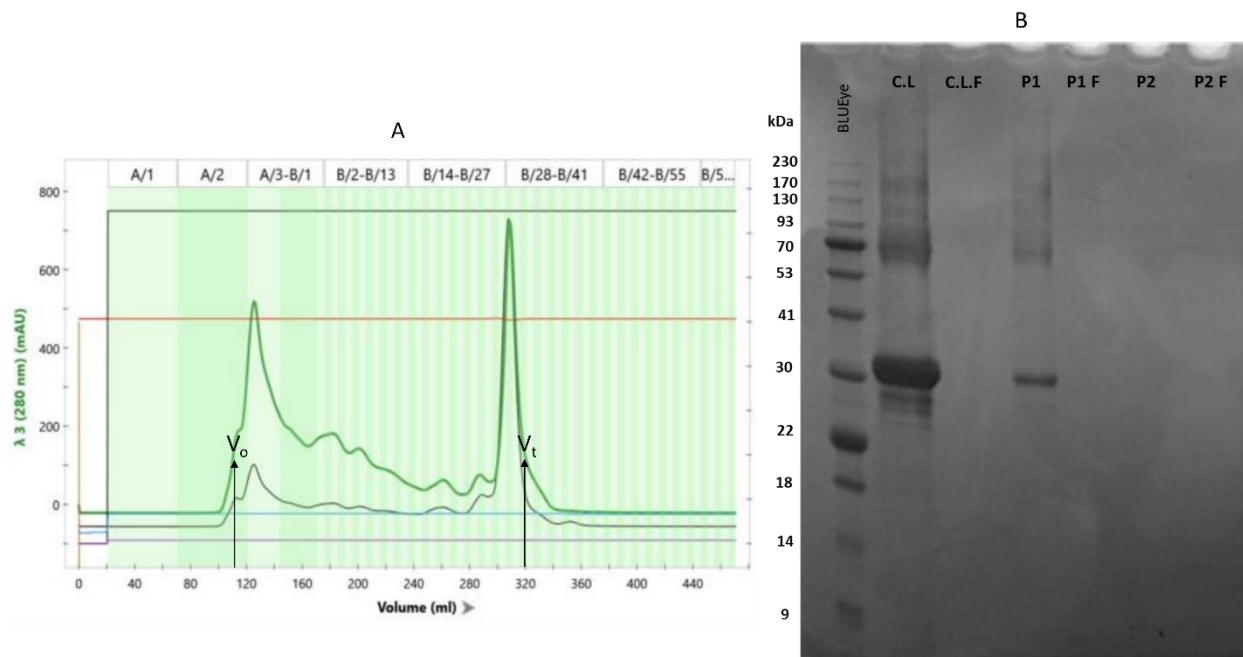


Figure 4.7 SEC purification of deGFP-BMP2. A: chromatograph with labels of V_o as void volume and V_t as total column volume. B: Non-reduced LDS PAGE of the second HisTrap column loading fraction (C.L), peak 1 (P1) and peak 2 (P2) of the SEC. F= Filtrate from the 10K MWCO Vivaspin 20 PES filters after concentrating the samples. BLUEeye as protein ladder.

In Figure 4.7 B proteins were only present in the concentrated samples, meaning the filter on the spin concentrator was able to retain them. No bands were visible from peak 2, which can be explained by the presence of imidazole, an aromatic compound present in the buffer which is too small to appear on LDS PAGE. This fraction was also not green under black light, whereas 'C.L' and 'Peak 1' were. Peak 1 lane appeared to consist of two major populations, one above 30 kDa and the other around 70 kDa which could be attributed to monomeric and dimeric forms of the chimera, respectively. These bands presented at a lower molecular weight than expected, but perhaps the lack of charge neutralization by SDS and therefore, inability to separate proteins based on weight alone could contribute to this difference. Figure 4.7 B also revealed that the column loading fraction during the IMAC purification had the highest protein concentration (1.248 mg/mL) when compared to the SEC peak 1 (0.242 mg/mL). Unfortunately, neither the IMAC nor SEC were sufficient methods for monomeric and dimeric separation.

GelAnalyzer 19.1 software (www.gelanalyzer.com) by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc, was used to measure the intensity of the bands and the ratio of monomer to dimer (Figure 4.8). In the C.L fraction (Figure 4.8 A), monomeric deGFP-BMP2 accounted for 59% and dimeric for 26% of the total lane (Figure 4.8 B), whereas in the peak 1 fraction (Figure 4.8 C), monomeric chimera accounted for 71% and dimeric for 25% (Figure 4.8 D). This finding suggests that functional protein recovery from refolding is rather inefficient, which can be observed throughout literature^{38, 39}. At the time, no further separation of monomer and dimer was undertaken due to limited time. However, it could be possible to further separate dimeric form from monomeric form by utilising a heparin-based affinity chromatography based on its interactions with growth factors including BMPs, or by using a longer column to allow for larger distance between the monomer and dimer⁴⁰. For protein characterisation experiments and cell culture, either C.L or P1 protein aliquots were used.

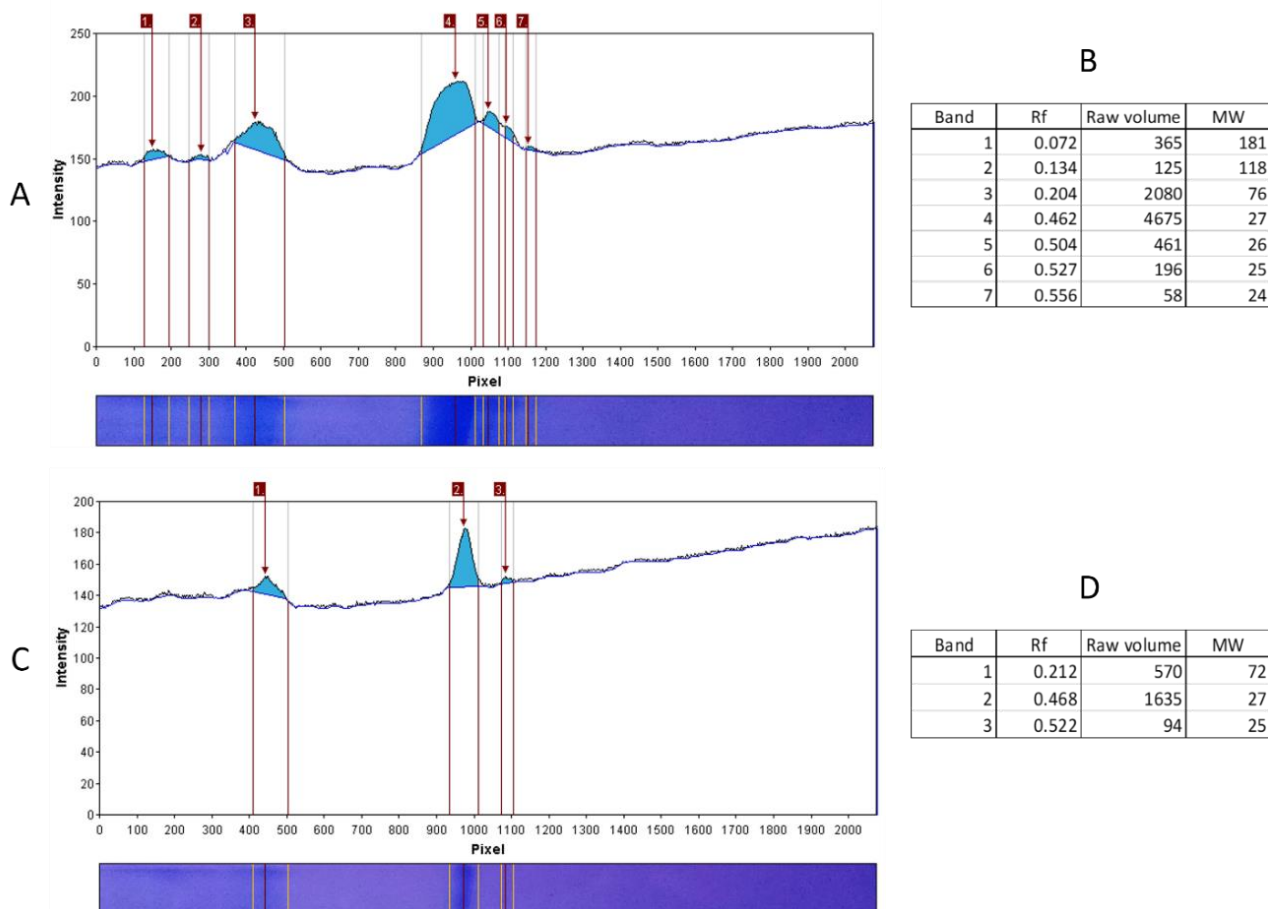


Figure 4.8 Lane content analysis from Figure 4.7 B SDS PAGE, performed by GelAnalyzer. A: Lane C.L with identified peaks in intensity. B: Values obtained from GelAnalyzer for lane C.L. C: Lane P1 with identified peaks in intensity. D: Values obtained from GelAnalyzer for lane P1.

4.3.2 SECONDARY STRUCTURE ANALYSIS OF CHIMERA

The secondary structure analysis of the deGFP-BMP2 was performed by circular dichroism (CD) spectroscopy and the data was deconvoluted using BeStSel software²⁶. EGFP, expressed and purified by Dr Ioatzin Rios de Anda (School of Physics, University of Bristol) was used alongside the chimera in the characterisation experiments. Although the secondary or tertiary structure of EGFP cannot be directly compared to that of the chimera, or even that of deGFP due to some deletions of the amino acid residues, it can serve as a useful model of a fluorophore. Currently, there are no structural data in literature available on deGFP-BMP2 since it is a novel fusion protein. The CD data obtained from EGFP and deGFP-BMP2 are presented in Figure 4.9.

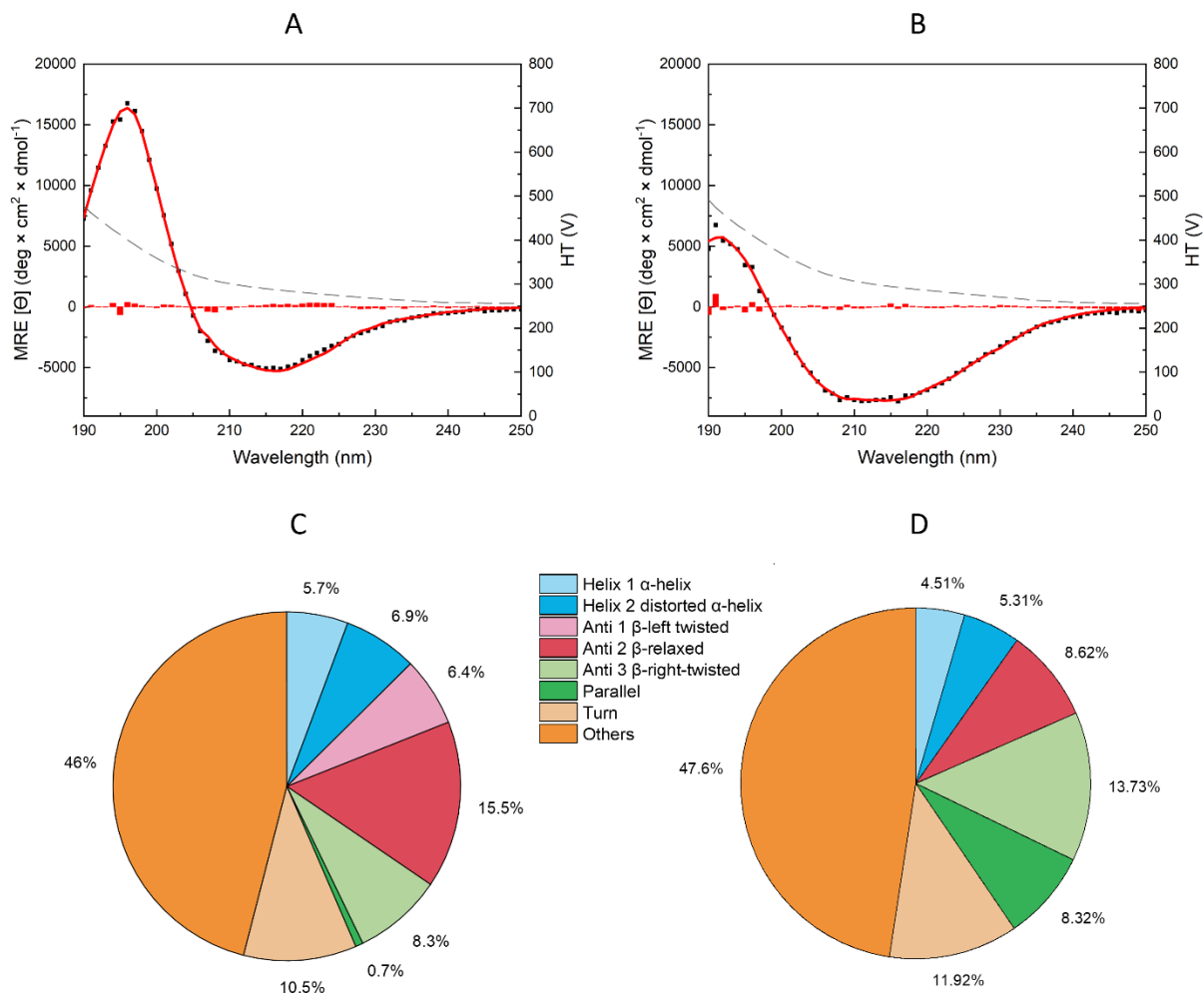


Figure 4.9 Circular dichroism (CD) spectra of A: EGFP and B: deGFP-BMP2 chimera. Data was deconvoluted with BeStSel software and displayed as fitted spectrum (red line), experimental data (black squares) and residuals (red columns), as well as the associated high-tension voltage (black dashes) displayed on the right y-axis. Pie chart of secondary structure composition of C: EGFP and D: deGFP-BMP2 chimera calculated from BeStSel software. ‘Others’ refer to protein structures consisting of 3_{10} helix, π -helix, β -bridge, bend, loop/irregular and invisible regions.

In a CD spectrum, a disordered protein can be identified by a negative peak around 195 nm and low ellipticity past 210 nm, whereas highly α -helical proteins have a positive peak around 193 nm and two negative dips at 208 nm and 222 nm²⁵. The CD spectra of both EGFP and deGFP-BMP2 (Figure 4.9 A and B respectively) show a positive peak in the 190-200 nm regions and negative peak in the 210-220 nm regions, with maxima and minima at 196/217 nm for EGFP and 191/216 nm for deGFP-BMP2. These spectral

features are typical of β -sheets which make up for a large proportion of the structure of both proteins. β -rich proteins show a broad range of spectra due to the variance in their backbone positioning caused by β -sheet twisting to different degrees⁴¹. The values of EGFP obtained here mostly agree with the values in literature where 13% of EGFP secondary structure was found to be helical, 47% comprised of β -sheets and the remaining 40% was attributed to loops of ‘others’⁵. BeStSel is a software which can distinguish between parallel and antiparallel β -sheets, and further categorising antiparallel into three conformations based on set boundaries of the antiparallel β -sheet twist angles. It was noted in Figure 4.9 C and D that in both proteins, antiparallel β -sheets were more prevalent than parallel β -sheets. This difference can be explained by a more favourable alignment of the dipoles due to closer packing in the antiparallel β -sheets and therefore, a more stable conformation⁴². The lowest secondary structure contribution from antiparallel β -sheets were those of a left-hand twist, with 6.4% for EGFP and <0.01% for deGFP-BMP2. In general, a left-hand twist sheet has a higher free energy than a right-hand twist, making it a less favourable conformation⁴³. Pie charts in Figure 4.9 C and D also revealed that the chimera has a lower content of α -helices (9.8%) than EGFP (12.6%), which could be contributed by the deletion of the small α -helix at the N-terminal of deGFP. Although the secondary structure values of BMP2 in literature point to a low α -helix content at 10.4% and high β -sheet content at 44.3%, which could intensify the largely β -sheet structure of the chimera as well⁴⁴. Generally, the chimera’s secondary structure consists of more stable conformations than those in EGFP, but at this stage it was difficult to determine whether the stability in the structure originated from BMP2 addition to the chimera or the deletions in the fluorophore to create deGFP. The structural integrity over a range of temperature of both proteins was analysed and can be seen in Figure 4.10. The chimera was expected to have a two-step (monomeric BMP2 with deGFP) or a three-step (dimeric BMP2 with deGFP) unfolding transition if all the subunits were to unfold

independently. Denaturing of the fluorophore would be characterised by a decrease in the β -sheet content and an increase in disorder.

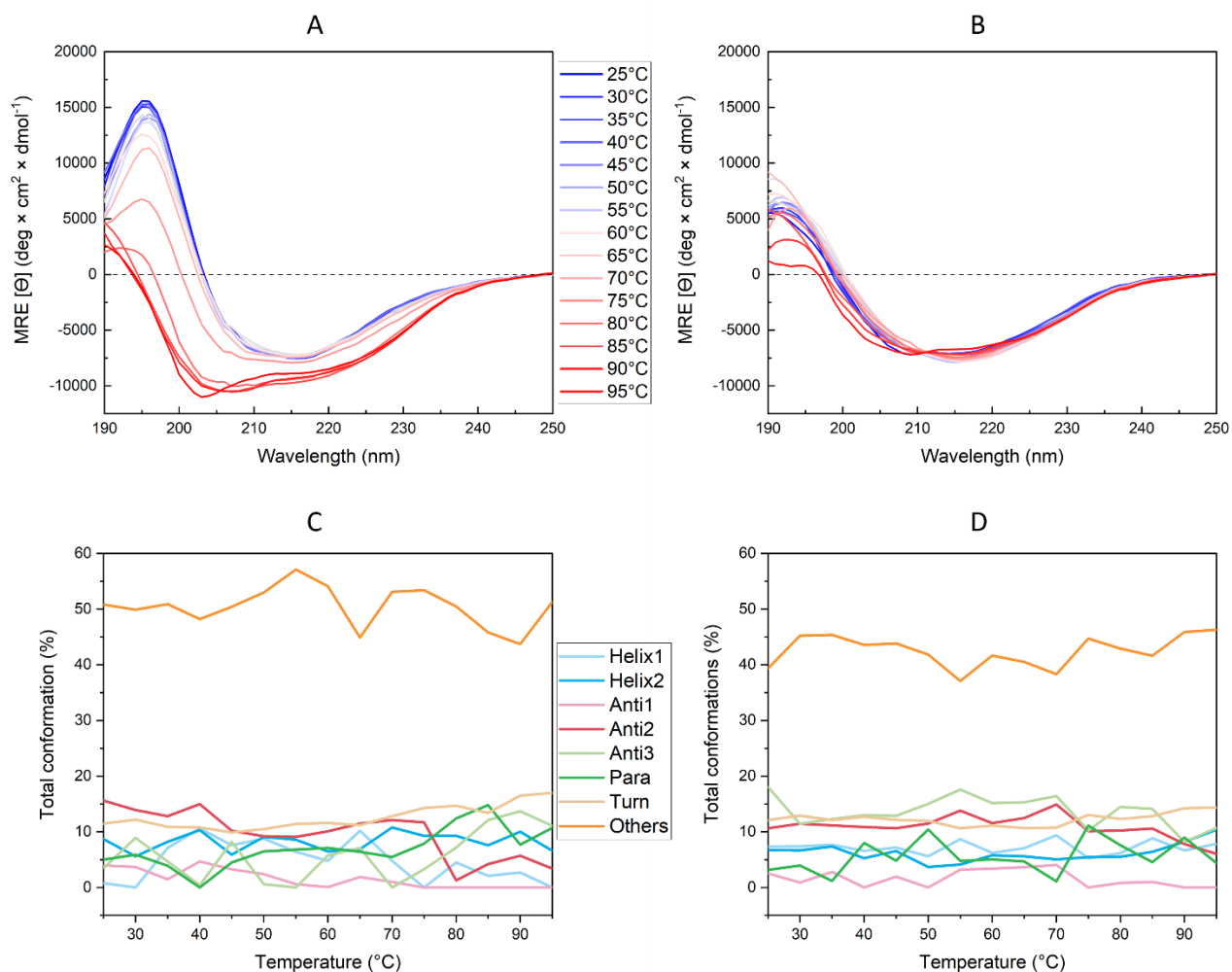


Figure 4.10 Fitted circular dichroism (CD) spectra obtained at a temperature gradient with A: EGFP and B: deGFP-BMP2 chimera. Total conformation percentages of secondary structure across the temperature gradient of C: EGFP and D: deGFP-BMP2 chimera. Values obtained from BeStSel deconvolution.

Contrary to this, the overall structure of the chimera did not appear to denature under the measured temperature conditions. To add to this, the secondary structure composition of the chimera did not have an observable change over the temperature ramp from 25°C to 95°C (Figure 4.10 D). Perhaps the fusion of the deGFP with BMP2 or the reduced amino acid residues of deGFP helped stabilise the overall structure of the chimera. The

EGFP thermal spectra in Figure 4.10 A was shifting towards low epsilon values with an increase in temperature, but a significant unfolding would flatten the β -sheet model over the measured wavelength towards $0 \text{ M}^{-1}\text{cm}^{-1}$. No significant changes in secondary structure composition of EGFP with an increase in temperature was found except for a decrease in relaxed β -sheet and a small increase in right-twisted β -sheet (Figure 4.10 C).

4.3.3 TERTIARY STRUCTURE ANALYSIS OF CHIMERA

The first technique used to aid characterisation of the tertiary structure of the chimera, alongside EGFP was UV-vis spectroscopy (Figure 4.11).

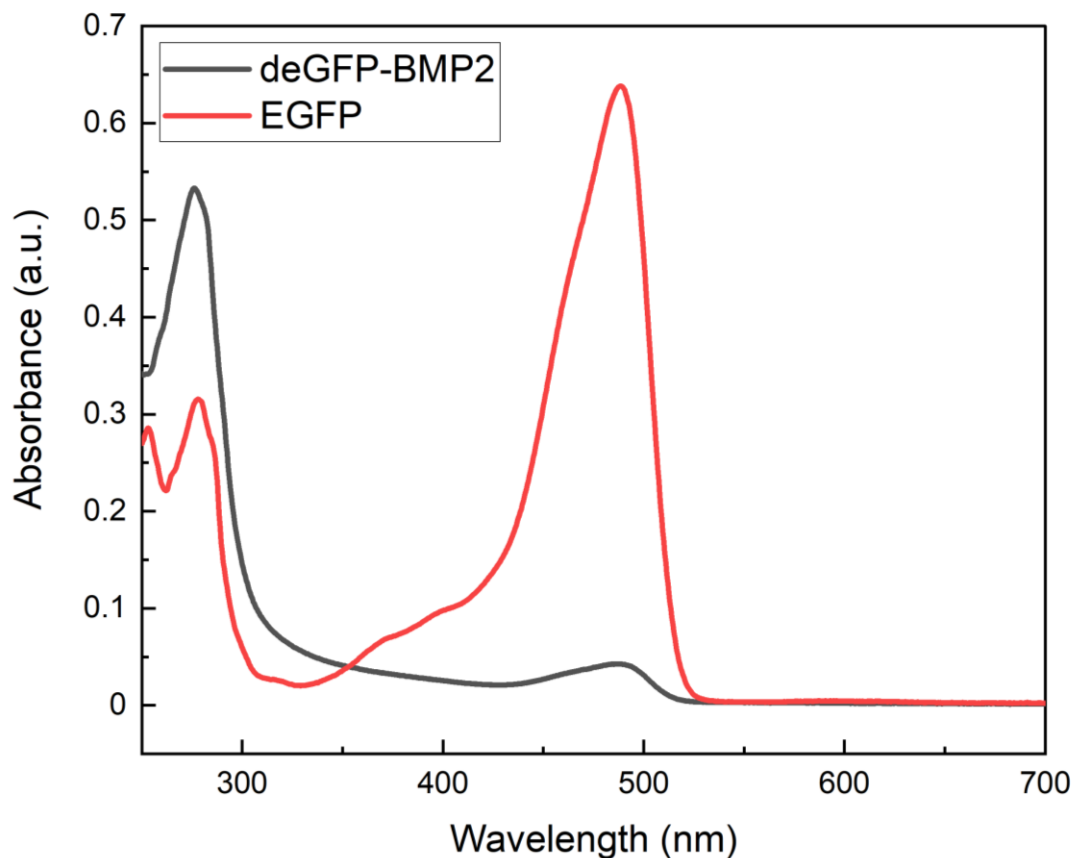


Figure 4.11 UV-vis spectrum of EGFP (red) and deGFP-BMP2 chimera (black). Measurements were taken at 25°C, in buffer of 20 mM Tris HCl, 150 mM NaCl, 4 M urea, pH 8.0.

Two absorbance peaks were predicted at 280 nm and 487 nm for both proteins. The peak at 280 nm corresponds to the absorbance of aromatic residues in both proteins, including tyrosine (Y) and tryptophan (W) amino acids, whereas the peak at 488 nm represents the

chromophore, specifically the threonine, tyrosine and glycine residues at locations 65, 66, 67 respectively, within the mature β -barrel^{45, 5}.

In Figure 4.11, both proteins give rise to these two expected peaks with neither of the peaks being red- or blue-shifted. However, the ratio of A_{487} to A_{280} maxima of the chimera is 0.11:1, whereas the ratio for EGFP is 2.03:1. The addition of the BMP2 sequence to the fluorescent protein did not incur an increase or a decrease of residues that would contribute to the absorbance at 487 nm, although the present tyrosine and tryptophan amino acids in BMP2 does influence the overall extinction coefficient of the chimera, which was theoretically calculated at $41,465 \text{ M}^{-1} \text{ cm}^{-1}$ and BSA-determined at $46,619 \text{ M}^{-1} \text{ cm}^{-1}$ with the absorbance at 280 nm. Whereas the extinction coefficient of EGFP is $55,000 \text{ M}^{-1} \text{ cm}^{-1}$ at 280 nm⁵. Consequently, the absorbance difference at 487 nm between EGFP and deGFP-BMP2 cannot be attributed to this. Additionally, the total protein detected at 280 nm might not be fully composed of protein fused with deGFP, which would manifest as a low $A_{487}:A_{280}$ ratio. This would have to be further delineated with techniques such as mass spectrometry for primary structure, small angle X-ray scattering for conformational tertiary structure in solution, or protein crystallisation and structure determination by synchrotron radiation for higher resolution. Another possible reason why the absorbance peak at 487 nm of the chimera was dwarfed by that of EGFP could be related to fluorescence quenching. The direct fusion of proteins or elements to the fluorophore has previously been shown to affect the tertiary structure of the fluorophore, contributing to quenching⁴⁶. To determine whether this was the case, fluorescence spectroscopy was conducted.

Within the chromophore of EGFP or deGFP, a reaction between the carboxyl carbon of residue 65 (threonine) and the amino nitrogen of residue 67 (glycine) occurs, generating imidazolin-5-one heterocyclic nitrogen ring. Further conjugation with tyrosine at location 66, results in the maturation of the chromophore⁵. When the chromophore absorbs light in the ultraviolet to blue range, green fluorescence is emitted at 511 nm. This excitation

and emission were measured for both the chimera and EGFP and presented in Figure 4.12.

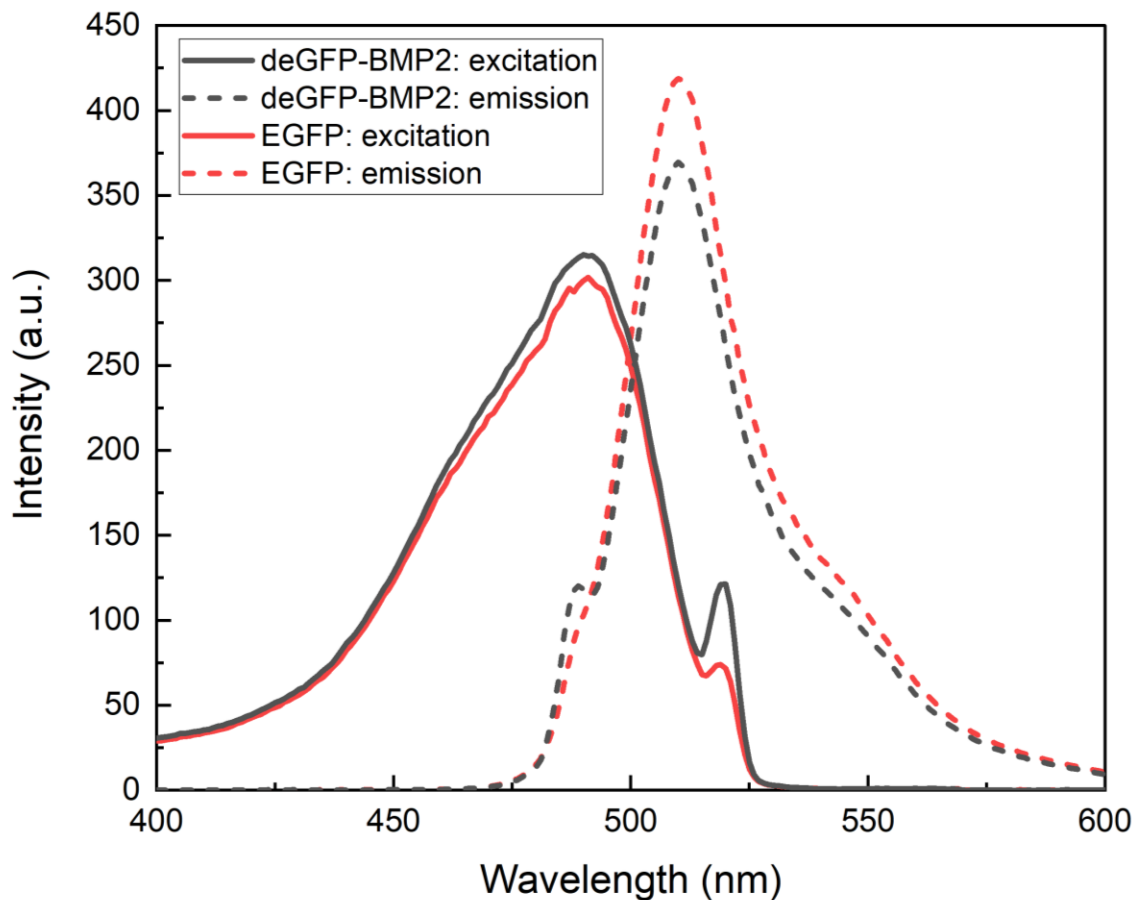


Figure 4.12 Fluorescence excitation (solid lines) and emission (dashed lines) spectrum of EGFP (red) and deGFP-BMP2 chimera (black). Measurements taken at 25°C in buffer of 20 mM Tris HCl, pH 8.0.

The excitation (487 nm) and emission (511 nm) peaks of the chimera and EGFP were consistent with the literature, and the shape of the spectra of EGFP is akin to that of chimera⁵. Therefore, the fluorescent properties of deGFP have not been significantly altered by the addition of the growth factor sequence and the fluorophore that is present has matured and folded correctly.

To further probe the tertiary structure of the chimera, dynamic light scattering experiments were performed to extract the hydrodynamic diameter size distribution (Figure 4.13).

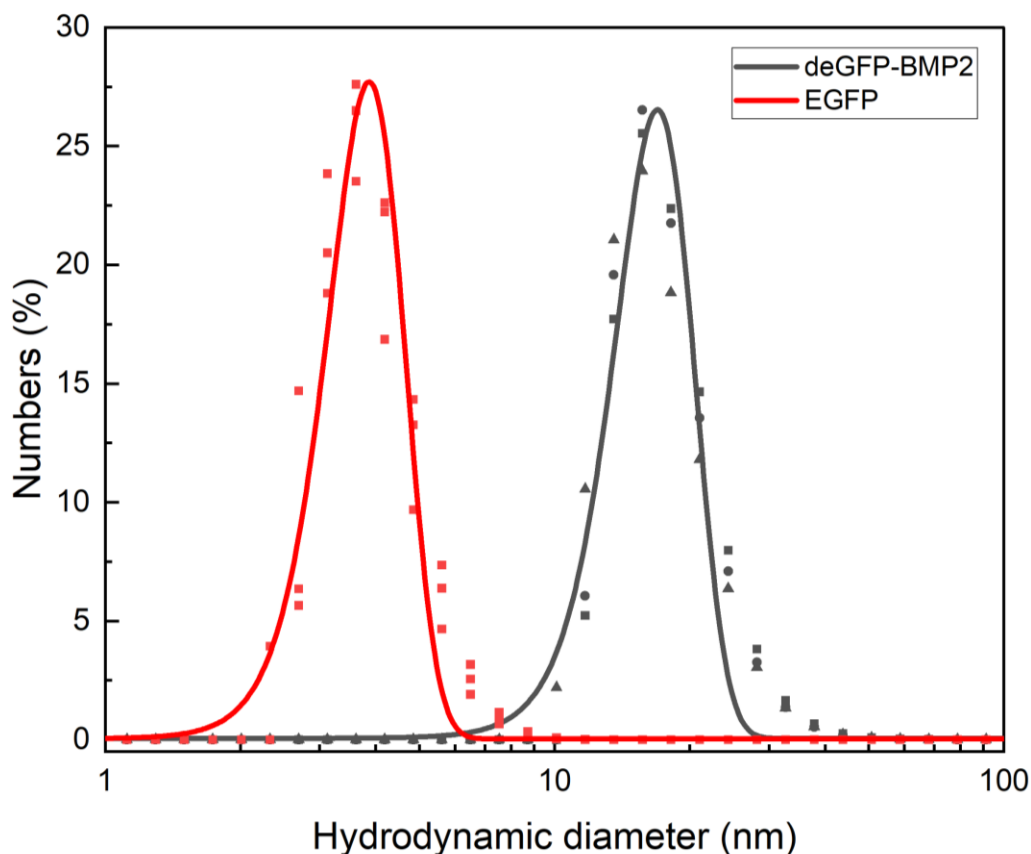


Figure 4.13 Dynamic light scattering (DLS) measurements to determine the hydrodynamic diameter of EGFP (red) and deGFP-BMP2 (black). The experimental data was fitted with a Gaussian curve (line plot). Three technical repeats.

Dynamic Light Scattering (DLS) is a technique that assumes the spherical nature of tested particles and the resulting outcome is quoted as hydrodynamic diameter, which is then fitted with a non-linear Gaussian function post-measurement. There is a myriad of proteins which do not fold into a spherical shape, therefore, DLS provides an approximate measurement based on this assumption. DLS is influenced by diffusion, which is relatively fast for small proteins, limiting resolution³². However, DLS can still provide useful information, especially when comparing differently sized proteins or aggregated states.

The size of the cylindrical EGFP has been resolved by X-ray diffraction and found to be 4.2 nm x 2.4 nm, and the hydrodynamic radius has been reported at 2.3 nm⁴⁷. Whereas, for a BMP2 dimer, the monomers assemble together into a ‘butterfly’ shape and as the crystal structure has been resolved at 2.7 Å by Scheufler *et al.*, the dimensions have been quoted at 7.0 nm x 3.5 nm x 3 nm¹⁰. The hydrodynamic radius of the BMP2 dimer has been noted in the range of 4-8 nm depending on the buffer conditions³⁷. The number of residues that were added to deGFP to create the chimera, including BMP2 sequence, His-tag, and a TEV cleavage site totals at 127 amino acids, a third of the total residues in the sequence. Each amino acid can contribute between 0.4 to 1 nm of length in a linear chain, and although the size can drastically change as the chain folds in 3D space, the 127 residues can significantly increase the size of the chimera, especially when BMP2 dimerises⁴⁸. Here, EGFP was found to have a hydrodynamic diameter of 3.87 ± 0.03 nm, whereas the diameter of chimera was calculated at 16.94 ± 0.13 nm. Significantly, there was no obvious evidence of aggregation in the sample. The size difference between the two proteins is consistent with the fact that the chimera has a longer amino acid sequence than EGFP. This large difference could be attributed to the dimerisation of the whole chimera, potentially with the two monomeric BMP2 bonded *via* disulphide bridges at the ‘palms of the hand’, each with a deGFP fluorophore attached at the ‘thumb’. This speculation aligns with the DLS data however, further studies should be carried out to confirm it, with X-ray crystallography being the most common method for studying protein structures, which can be coupled with small angle x-ray scattering (SAXS). It is also important to note that the DLS sample contained the mixture of monomeric and dimeric forms of the chimera due to the inability to separate at the size exclusion chromatography step of purification. For a more accurate DLS measurement, it would be crucial to have the distinct populations of the monomeric and dimeric deGFP-BMP2 separated.

4.4 CONCLUSIONS AND FURTHER WORK

The chimera, deGFP-BMP2, which was developed to be expressed using the in-house cell-free system, was also subjected to expression and purification from *E.coli* cells to provide quantities sufficient for structure characterisation, as well as stem cells differentiation studies described in Chapter 5. From the small-scale chimera expression experiments, it was evident that the overexpressed protein was present in the insoluble fraction, which had to be solubilised and refolded to improve the formation of disulphide bonds in the dimeric chimera. There are numerous bacterial strains that could have been explored as means to express recombinant deGFP-BMP2, including Rosetta-gami, a strain based off the Rosetta used in this project, however, this particular strain also aids in the correct folding of proteins with disulphide bonds⁴⁹. Perhaps the use of the Rosetta-gami strain would suffice in dimerisation efficiency, negating the need for further refolding as had to be performed here. The separation of the monomeric and dimeric forms using nickel IMAC or SEC proved too challenging and the final fractions contained both conformations. In future, the heparin binding ability of BMP2 could be utilised for an improved separation by using Heparin Sepharose 6 Fast Flow column with a gradient elution⁵⁰. Besides the synthesis and purification of the chimera, the characterisation using CD, DLS, UV-Vis and fluorescence spectroscopy was also discussed in this chapter.

EGFP was used here to aid comparisons with the chimera for these techniques, however, more appropriate controls would include deGFP as a fluorophore and unfused BMP2. At the time, expression of deGFP was attempted but did not yield a significant amount. Circular dichroism performed on the chimera and EGFP showed that the secondary structure corresponds to the predicted high β -sheet content for both proteins and the structure is not significantly affected by temperature. Although the use of benchtop CD yielded novel characterisation, the use of synchrotron radiation CD with much greater photon flux would produce higher resolution results with improved signal to noise ratio. The tertiary structure analysis performed demonstrated that by fusing the BMP2 sequence

along with affinity and cleavage tags to the fluorophore did not impede the fluorescence properties, but the ratio of growth factor to the fluorophore would need advanced delineation. DLS results have shown the hydrodynamic radius of EGFP and deGFP-BMP2 to be estimated at 3.87 ± 0.03 nm, and 16.94 ± 0.13 nm respectively. This hydrodynamic radius difference can be attributed to the larger size of the chimera when the BMP2 dimerization generates two BMP2 monomers and two deGFP molecules. There are other methods such as small x-ray scattering, 2D NMR spectroscopy or electron microscopy that would provide further understanding of the protein physical structure. Matrix assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS) could also have helped measure an accurate mass of the chimera and improve the understanding of the composition of the protein⁵¹.

4.5 BIBLIOGRAPHY

1. Yerneni, S. S., Adamik, J., Weiss, L. E. & Campbell, P. G. Cell trafficking and regulation of osteoblastogenesis by extracellular vesicle associated bone morphogenetic protein 2. *Journal of Extracellular Vesicles* 10, e12155 (2021).
2. Alborzina, H. *et al.* Quantitative kinetics analysis of BMP2 uptake into cells and its modulation by BMP antagonists. *Journal of Cell Science* 126, 117–127 (2012).
3. Kang, W., Lee, D. S. & Jang, J. H. Evaluation of Sustained BMP-2 Release Profiles Using a Novel Fluorescence-Based Retention Assay. *PLOS ONE* 10, e0123402 (2015).
4. Yang, F., Moss1, L. G. & Phillips, G. N. The molecular structure of green fluorescent protein. *Nature* 14, 1246–1251 (1996).
5. Arpino, J. A. J., Rizkallah, P. J. & Jones, D. D. Crystal Structure of Enhanced Green Fluorescent Protein to 1.35 Å Resolution Reveals Alternative Conformations for Glu222. *PLoS ONE* 7, 47132 (2012).
6. Arpino, J. A. J., Rizkallah, P. J. & Jones, D. D. Structural and dynamic changes associated with beneficial engineered single-amino-acid deletion mutations in enhanced green fluorescent protein. *Acta Crystallogr D Biol Crystallogr* 70, 2152–2162 (2014).
7. Li, X. *et al.* Deletions of the *Aequorea victoria* Green Fluorescent Protein Define the Minimal Domain Required for Fluorescence. *Journal of Biological Chemistry* 272, 28545–28549 (1997).
8. Liu, S.-S. *et al.* Structural plasticity of green fluorescent protein to amino acid deletions and fluorescence rescue by folding-enhancing mutations. *BMC Biochemistry* 16, (2015).

9. Rengachary, S. Bone morphogenetic proteins: basic concepts. *Neurosurg. Focus* 13, (2002).
10. Scheufler, C., Sebald, W. & Hülsmeier, M. Crystal structure of human bone morphogenetic protein-2 at 2.7 Å resolution. *Journal of Molecular Biology* 287, 103–115 (1999).
11. Rider, C., Holloway, R. & Mulloy, B. Bone morphogenetic protein and growth differentiation factor cytokine families and their protein antagonists Analysis View project. *Biochemical Journal* 429, 1–12 (2010).
12. Mueller, T. D. & Nickel, J. Promiscuity and specificity in BMP receptor activation. *FEBS Letters* 586, 1846–1859 (2012).
13. Everson, R. J. & Parker, H. E. Zinc binding and synthesis of eight-hydroxy-quinoline-agarose. *Bioinorganic Chemistry* 4, 15–20 (1974).
14. Porath, J., Carlsson, J., Olsson, I. & Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258, 598–599 (1975).
15. Gaberc-Porekar, V. & Menart, V. Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods* 49, 335–360 (2001).
16. Hochuli, E., Dobeli, H. & Schac, A. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. *Journal of Chromatography* 411, 177–184 (1987).
17. Barth, H. G., Boyes, B. E. & Jackson, C. Size Exclusion Chromatography. *Analytical Chemistry* 68, 445–466 (1996).

18. Lathe, G. & Ruthven, C. The separation of substances on the basis of their molecular weights, using columns of starch and water. *Biochem. J* 60, 34–34 (1955).
19. Štulík, K., Pacáková, V. & Tichá, M. Some potentialities and drawbacks of contemporary size-exclusion chromatography. *Journal of Biochemical and Biophysical Methods* vol. 56 1–13 Preprint at [https://doi.org/10.1016/S0165-022X\(03\)00053-8](https://doi.org/10.1016/S0165-022X(03)00053-8) (2003).
20. Srinivas, P. R. Introduction to Protein Electrophoresis. *Methods in Molecular Biology* 869, 23–28 (2012).
21. Shapiro, A. L., Viñuela, E. & v. Maizel Jr., J. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochemical and Biophysical Research Communications* 28, 815–820 (1967).
22. Walker, J. M. Nondenaturing Polyacrylamide Gel Electrophoresis of Proteins. *The Protein Protocols Handbook* 57–60 (2002) doi:10.1385/1-59259-169-8:57.
23. Micsonai, A. *et al.* Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc Natl Acad Sci U S A* 112, E3095–E3103 (2015).
24. Rodger, A. & Marshall, D. Beginners guide to circular dichroism. *The Biochemist* 43, 58–64 (2021).
25. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* 1, 2876–2890 (2006).
26. Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Research* 46, 315–322 (2018).

27. Weckhuysen, B. M. In-situ spectroscopy of catalysts. in *In-situ Spectroscopy of Catalysts* 255–270 (American Scientific Publishers, 2004).
28. Akash, M. S. H. & Rehman, K. Ultraviolet-Visible (UV-VIS) Spectroscopy. *Essentials of Pharmaceutical Analysis* 29–56 (2020) doi:10.1007/978-981-15-1547-7_3.
29. Brand, L. & Johnson, M. L. Methods in enzymology, Fluorescence Spectroscopy. in vol. 450 (Elsevier, 2008).
30. Lakowicz, J. R. *Principles of fluorescence spectroscopy*. (Springer, 1999).
31. Stetefeld, J., Mckenna, S. A. & Patel, T. R. Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophysical Reviews* 8, 409–427 (2016).
32. Stokes, G. G. On the Theories of the Internal Friction of Fluids in Motion and of the Equilibrium and Motion of Elastic Solids. *Transactions of the Cambridge Philosophical Society* 8, 287–305 (1845).
33. Kane, J. F. Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli. *Current Opinion in Biotechnology* 6, 494–500 (1995).
34. Lobstein, J. *et al.* SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. *Microbial Cell Factories* 11, 56 (2012).
35. Paliy, O. & Gunasekera, T. S. Growth of E. coli BL21 in minimal media with different gluconeogenic carbon sources and salt contents. *Applied Microbial and Cell Physiology* 73, 1169–1172 (2007).
36. Wang, E. A. *et al.* Recombinant human bone morphogenetic protein induces bone formation. *Proc Natl Acad Sci U S A* 87, 2220–2224 (1990).

37. Heinks, T. *et al.* Optimized expression and purification of a soluble BMP2 variant based on in-silico design. *Protein Expression and Purification* 186, 105918 (2021).
38. Eiberle, M. & Jungbauer, A. Technical refolding of proteins: Do we have freedom to operate? *Biotechnology Journal* 5, (2010).
39. Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, M. & Kumar Panda, A. Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories* 14, (2015).
40. Bolten, S. N., Rinas, U. & Scheper, T. Heparin: role in protein purification and substitution with animal-component free material. *Applied Microbiology and Biotechnology* 102, 8647 (2018).
41. Sreerama, N. & Woody, R. W. Structural composition of β I- and β II-proteins. *Protein Science* 12, 384 (2003).
42. Chou, K. C., Pottle, M., Némethy, G., Ueda, yuzo & Scheraga, H. A. Structure of β -sheets. Origin of the right-handed twist and of the increased stability of antiparallel over parallel sheets. *Journal of Molecular Biology* 162, 89–112 (1982).
43. Chothia, C. Conformation of twisted β -pleated sheets in proteins. *Journal of Molecular Biology* 75, 295–302 (1973).
44. Marquetti, I. & Desai, S. Orientation effects on the nanoscale adsorption behavior of bone morphogenetic protein-2 on hydrophilic silicon dioxide. *RCS Advances* 9, (2019).
45. Gill, S. & Hippel, P. H. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Analytical Biochemistry* 182, 319–326 (1989).

46. Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* 6, 178–182 (1996).
47. Hink, M. A. *et al.* Structural Dynamics of Green Fluorescent Protein Alone and Fused with a Single Chain Fv Protein . *Journal of Biological Chemistry* 275, 17556–17560 (2000).
48. Ching, C. B., Hidajat, K., Uddin, M. S. & Ridge, K. Kinetic Parameters of Smaller Molecular Size Amino Acids on KX Zeolite Crystals via Liquid Chromatographic Techniques. *Separation Science and Technology* 24, 581–597 (1989).
49. Zarkar, N., Nasiri Khalili, M. A., Khodadadi, S., Zeinoddini, M. & Ahmadpour, F. Expression and purification of soluble and functional fusion protein DAB389IL-2 into the E. coli strain Rosetta-gami (DE3). *Biotechnology and Applied Biochemistry* 67, 206–212 (2020).
50. Gandhi, N. S. & Mancera, R. L. Prediction of heparin binding sites in bone morphogenetic proteins (BMPs). *Biochimica et Biophysica Acta - Proteins and Proteomics* 1824, 1374–1381 (2012).
51. Webster, J. & Oxley, D. Protein identification by MALDI-TOF mass spectrometry. in *Methods in Molecular Biology* vol. 800 227–240 (Humana Press, 2012).

CHAPTER 5.
DIFFERENTIATION OF HUMAN
MESENCHYMAL STEM CELLS
USING DEGFP-BMP2

5.1 INTRODUCTION

Bone morphogenetic proteins (BMPs) play a major role throughout the process of endochondral bone formation and fracture healing. In both instances, the initial phase, directed by BMPs, is the recruitment of mesenchymal stem cells and their proliferation and differentiation towards the chondroblast and osteoblast lineages¹. BMP2, 6 and 9 are the major inducers of human mesenchymal stem cells (hMSCs) differentiation². BMP2 transduces signalling by binding to the heteromeric BMP type I and II receptors (BMPRI and BMPRII), which are transmembrane serine/threonine kinases (Figure 5.1)³. BMPRI is subdivided into activin-like kinase 3 (ALK3/BMPRIA), activin-like kinase 6 (ALK6/BMPRIIB) and activin-like kinase 2 (ALK2/ACVRI). In fact, the binding occurs between BMP2 and type I receptor which then heterodimerises with the type II receptor; the constitutively active type II receptor transphosphorylates the glycine/serine-rich region of type I kinase receptor, thereby activating it⁴. The active BMPRI switches on Smad proteins 1, 5 and 8 by phosphorylation of the Ser-Ser-Val/Met-Ser motif, followed by the formation of heteromeric complexes with Smad4. Once the complex is completed, it becomes translocated to the nucleus where it remodels the chromatin to regulate the transcription of genes involved in osteogenesis, with runt-related transcription factor 2 (*RUNX2*) being the master gene. There are also non-canonical pathways of intracellular signalling initiated by BMP2 binding to the receptors, such as MKK-p38 MAPK or MKK-ERK1/2 cascades⁵. It has been shown that in both cases of Smad-dependent and independent signalling pathways, BMPRI and BMPRII are localised in caveolae and clathrin-coated pits, and the initiation of the pathways occur in those domains⁶. The process of osteogenic differentiation continues with the upregulation of the essential transcription factor, *RUNX2*. Mice studies with deleted *RUNX2* have shown to be deficient in bone formation due to the arrested osteoblast maturation⁷. This transcription factor, along with osterix (*OSX*) and drosophila distal-less 5 (*DLX5*), regulates the

expression of several osteogenic genes that commit the osteo-chondroprogenitor cell into a pre-osteoblast⁸.

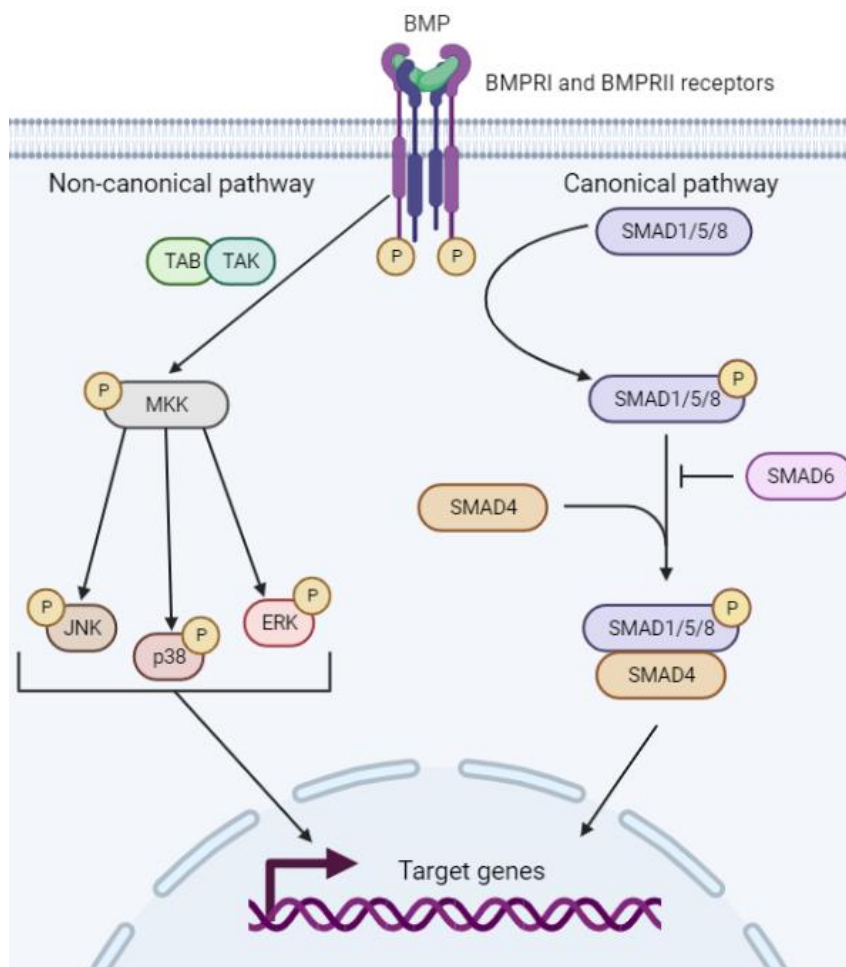


Figure 5.1 Schematic representation of the BMP2 signalling cascades including the canonical Smad and non-canonical pathways. Figure created in biorender.com.

Some of the osteogenic genes that are upregulated at this stage include collagen type I alpha 1 (*Col1A1*) and alkaline phosphatase (*ALP*), both of which continue to be expressed in the later stages of osteoblast differentiation. Once mature, osteoblasts secrete the extracellular matrix (ECM) and whilst maturing, the matrix becomes mineralised by the activity of osteoblastogenic markers including osteocalcin (*OCN/BGLAP*), osteopontin (*OPN*) and bone sialoprotein (*BSP*), which are promoters of calcium deposition, bone mineralisation and hydroxyapatite crystal formation, respectively⁹. At this stage, after

new bone deposition has been completed, osteoblasts can either transform into osteocytes, convert into bone lining cells or they can undergo apoptosis.

In this chapter, deGFP-BMP2 produced from either *E.coli* or the CFES are investigated through their interactions with hMSCs, and the impact on osteogenic differentiation. The assays performed here are important to distinguish whether CFES-made novel chimera has the osteogenic potential comparable to *E.coli*-derived and recombinant human BMP2, as well as to examine the impact CFES components may have on hMSCs culture. Initially, the binding of deGFP-BMP2 to hMSCs is probed using confocal microscopy and flow cytometry. The phosphorylation of Smad 1/5/8 as an effect of active BMP2 binding to BMPRI/BMPRII is also investigated using Western blotting. Finally, deGFP-BMP2-induced osteogenic differentiation of hMSCs is verified by monitoring alkaline phosphatase activity, quantitative polymerase chain reaction (qPCR) of *RUNX2*, *Col1A1*, *ALP*, *BGLAP* genes, and matrix maturation and mineralisation through Alizarin Red S staining.

5.2 MATERIALS AND METHODS

5.2.1 CELL CULTURE

Human mesenchymal stem cells (hMSCs) were obtained from Bristol Southmead Hospital in accordance with their Research Ethics Committee guidelines. Frozen samples (1×10^6 /mL cells) were retrieved from liquid nitrogen storage and were cultured as a monolayer in T175 flasks with 25 mL expansion media consisting of low-glucose Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 1% antibiotics (penicillin and streptomycin), 10% foetal bovine serum (FBS), 1% Glutamax and 1 ng/mL human fibroblast growth factor (FGF). Cell handling was carried out in SAFE 2020 laminar flow hoods (Thermo Fisher Scientific), whereas cell incubation has taken place in Hera Cell 150 incubator (Thermo Fisher Scientific) at 37°C and 5% CO₂ atmosphere. The supplemented DMEM was replaced with fresh media every two days until confluency of 90-95% has reached. The cells were then either passaged or used in subsequent experiments. To achieve this, the cells had to be detached from the T175 flask. The media was removed from the flask and the cell monolayer was washed twice with 15 mL of phosphate buffered saline (PBS). The cell monolayer was then submerged in 6 mL of trypsin/EDTA (Sigma Aldrich) and the flask was placed in a Hera Cell 150 incubator for 5 minutes to activate trypsin. After the incubation, the flask was gently tapped to ensure full cell detachment and 12 mL of DMEM with FBS was added to prevent further digestion by trypsin. The cell suspension was transferred into a 50 mL falcon tube and centrifuged at 1500 rpm for 5 minutes using Sorvall Legend RT centrifuge (Thermo Scientific). The resulting cell pellet was carefully resuspended in 1-5 mL fresh complete media by pipetting and gentle falcon tube tapping. Cell counting was performed by adding 10 µL of the cell suspension to the chambers of Neubauer haemocytometer (Hawksley) and manually counted, taking any dilution factors into consideration. Once the cell number was known, the cells were distributed to new flasks or well-plates accordingly.

The most common cell culture media used in the assays outlined in this chapter are detailed in Table 5.1.

Abbreviation	Media name	Components
S.M.	Starvation media	Low glucose Dulbecco's Modified Eagle's Medium (DMEM) (Merck Life Science)
E.M.	Expansion media	Low glucose Dulbecco's Modified Eagle's Medium (DMEM) (Merck Life Science) 1% penicillin and streptomycin (Merck Life Science) 10% fetal bovine serum (FBS) (Merck Life Science) 1% Glutamax (Life Technologies) 1 ng/mL human fibroblast growth factor (FGF) (Peprotech)
O.M.	Osteogenic media	Low glucose Dulbecco's Modified Eagle's Medium (DMEM) (Merck Life Science) 1% penicillin and streptomycin (Merck Life Science) 10% fetal bovine serum (FBS) (Merck Life Science) 1% Glutamax (Life Technologies) 10 nM dexamethasone (Sigma Aldrich) 0.2 mM Ascorbic acid 2-phosphate (Sigma Aldrich) 10 mM β -glycerophosphate (Sigma Aldrich)

Table 5.1 Components of cell culture media used in the assays detailed below.

In the assays where CFES was used in cell culture media, it refers to a whole reaction volume with the extract, buffer solution and input DNA ran to completion (post 16 hours) with the confirmation of successful expression by fluorescence monitoring. Afterwards, the concentration of the expressed protein was estimated and the completed CFES reactions was filter sterilised by 0.22 μ m syringe filter (Sartorius) before distribution to individual wells along with the appropriate cell culture media.

5.2.2 CONFOCAL MICROSCOPY

Confocal microscopy is an optical imaging technique that allows for collection of multispectral images at high speed and high resolution¹⁰. In a typical setup, there are two pinhole apertures; a light source pinhole and a detector pinhole. The excitation light from a laser passes through the first pinhole and is reflected by a dichroic mirror to generate a point of illumination within the sample at a single focal plane. The fluorescence emitted from the same focal plane in the specimen passes through the dichromatic mirror and *via*

the second, detector pinhole into the photomultiplier detector. The detector pinhole ensures the removal of out-of-focus fluorescence which is a contributor to background noise. By acquiring multiple images at thin sections of the specimen, a 3-D reconstruction can be generated¹¹.

Human mesenchymal stem cells were seeded in 35 mm confocal dishes at 250,000 cells/dish with expansion media and incubated in Hera Cell 150 incubator for overnight adherence. Hoechst 33342 (Thermo Fisher) was diluted 1:2000 and CellMask Orange (Life technologies) was diluted 1:1000 with phenol-free DMEM. The staining dyes were added to the confocal dishes for 10 minutes and incubated at 37°C, 5% CO₂ atmosphere. The staining solution was then removed and the cells were washed three times in PBS before placing 400 µg/mL *E.coli*-produced deGFP-BMP2 chimera in FluoroBrite DMEM (Thermo Fisher) onto the seeded cells. The cells were imaged in intervals using a Leica SP8 AOBS confocal laser scanning microscope attached to a Leica DMI8 inverted epifluorescence microscope.

5.2.3 FLOW CYTOMETRY

Flow cytometry allows for analysis and sorting of single cells as they pass as a single-file stream in solution by multiple lasers. Each individual cell is detected for fluorescence as well as measured for visible light scatter, which scans in forward direction (FSC) and at 90° side scatter (SSC). FSC measures for the relative size of the cell, whereas SSC detects degree of granularity within the cell. The presence of fluorescence detectors allows for sorting of cells populations depending on their dyes or fluorescent reporters¹².

Human mesenchymal stem cells were seeded as 150,000 cells/well in a 24-well plate, flat bottom format. Once the cells settled at the bottom of the wells, the media was replaced with one of the following conditions: untreated expansion media and expansion media supplemented with 0.1 mg/mL EGFP, 0.1 mg/mL *E.coli*-produced deGFP-BMP2 chimera, completed deGFP-BMP2 CFES reaction with the final estimated concentration

of deGFP-BMP2 at 3.54×10^{-3} mg/mL and the completed deGFP CFES reactions with the final estimated concentration of deGFP at 3.54×10^{-3} mg/mL. The incubation with the stated conditions took place for 6 hours at 37°C and 5% CO₂ atmosphere. Afterwards, the media conditions were removed, and the cells were washed once in PBS before adding 100 µL/well of trypsin. Incubation of 5 minutes in the Hera Cell 150 incubator was sufficient for cell detachment, after which the cells were resuspended in 200 µL/well of PBS and the cell suspensions were transferred to individual 12 mm diameter round-bottom polystyrene flow cytometry tubes (Corning) and kept on ice until analysis. Measurements were taken on BD LSRFortessa X20 cell analyzer (BD Biosciences) using the BD FACS Diva software (BD Biosciences), recording 10,000 events per condition. The flow cytometry results were analyzed using FlowJo™ v10.8 Software (BD Life Sciences). Untreated cells were used for gating of hMSCs population with forward scatter area (FSC-A) vs side scatter area (SSC-A). Single cells were gated by FSC-A vs forward scatter height (FSC-H). The gating for GFP fluorescence was selected by FSC-A vs FITC-A:GFP.

5.2.4 WESTERN BLOTTING

Western Blotting is technique coupled to gel electrophoresis which allows for detection of specific proteins with the use of antibodies¹³. The separated proteins from SDS PAGE are transferred onto a nitrocellulose or polyvinylidene fluoride (PVDF) membrane using an electrical current. The solid support of the membrane enables the binding of antibodies specific to the protein of interest. By introducing antibodies with detection labels such as horseradish peroxidase (HRP) the location of the protein of interest can be easily determined.

Human mesenchymal stem cells were seeded as 250,000 cells/well in a 6-well plate (Corning) with expansion media, and incubated at 37°C and 5% CO₂ atmosphere until the cells attached to the bottom of the wells. The media was then changed to starvation media, and incubated overnight in Hera Cell 150 incubator. Six conditions were tested: negative control of starvation media only, positive control of 25 ng/mL commercial

rhBMP2, 25 ng/mL and 200 ng/mL *E.coli*-produced deGFP-BMP2 chimera, 10 μ L completed deGFP-BMP2 CFES reaction in 1 mL media, and 10 μ L completed deGFP CFES reaction in 1 mL media. Cells were incubated with the stated conditions for 4 hours at 37°C, 5% CO₂ atmosphere. Afterwards, the cells were washed with PBS and 200 μ L/well of RIPA lysis buffer with a protease inhibitor cocktail (Thermo Fisher) was added. The well-plate was then incubated on ice for 20 minutes before using 16 cm cell scrapers (Sarstedt) to detach cells from the bottom. Cells were further homogenised by passing the sample several times with a 21 G syringe needle. The samples were then collected in 1.5 mL Eppendorf tubes and centrifuged at 13,000 rpm, 4°C for 10 minutes in Eppendorf 5424R centrifuge. The collected supernatant was used as samples for SDS PAGE following the previously described protocol (Chapter 4, Section 4.2.4). The resulting polyacrylamide gel was applied to an activated PVDF membrane, sandwiched by two pieces of 0.2 μ m 8.3 cm x 7.3 cm filter paper along with the sponge pads. The PVDF membrane was activated by submersion in 100% methanol for a few minutes before equilibration in 1x Novex Tris-glycine Transfer Buffer (TrB) (Thermo Fisher). The transfer sandwich was assembled according to the Novex Mini Blot Modules instructions (Thermo Fisher) and the transfer occurred at 20 V, 180 mA, for 1 hour. The PVDF membrane was incubated at room temperature for 2 hours in Blocking buffer (Tris buffered saline (TBS), 0.1% Tween 20, 1% milk) on a rocker. The membrane was then incubated in 10 mL of primary antibody solution (TBS, 0.1% Tween 20, 1% milk, 10 μ L of anti-Phospho-Smad 1/5 Ser463/465, 41D10 rabbit mAb (1:1000) (Cell Signaling Technology)) at 4°C overnight. Four washes of 10 minutes were then performed using TBS with 0.1% Tween 20. The secondary antibody solution (TBS, 0.1% Tween 20, 1% milk, 10 μ L of Goat anti rabbit HRP conjugated IgG2b (1:1000) (Thermo Fisher)) was added to the membrane and incubated at room temperature for 1 hour. The membrane was washed four times with 20 mL of TBS with 0.1% Tween 20 with the final wash of just TBS. Pierce Western ECL Blotting Substrate (Thermo Fisher) was used according

to the instructions and the membrane was imaged on Amersham imager 600 (GE Healthcare). After imaging the blot, several washes in TBST were performed before adding a loading control antibody solution (TBS, 0.1% Tween 20, 1% milk, 10 μ L of rabbit anti-vinculin polyclonal PA1781 (1:1000 dilution) (Insight Biotechnology Limited). This loading control antibody solution was incubated on the membrane overnight at 4°C, then the membrane was washed in TBST four times before adding the secondary antibody solution (TBS, 0.1% Tween 20, 1% milk, 10 μ L of Goat anti rabbit HRP conjugated IgG2b (1:1000) (Thermo Fisher)) for 1 hour incubation at room temperature. PVDF membrane underwent six washes in TBST before being submerged in the ECL Western substrate for final imaging.

5.2.5 ALKALINE PHOSPHATASE ASSAY

Alkaline phosphatase (ALP) is an enzyme that catalyses the hydrolysis of phosphate esters at basic pH environments. Increased ALP levels are associated with active bone formation due to ALP being the by-product of osteoclast activity. Therefore, ALP is often regarded as an early marker of osteogenesis¹⁴. In the fluorometric assay, the ALP enzyme cleaves the phosphate group of 4-methylumbelliferyl phosphate disodium salt (MUP), generating a blue, fluorescent product, 7-hydroxy-4-methyl-coumarin¹⁵. This enables an easy spectrofluorimetric detection and quantification by producing a standard curve.

Human mesenchymal stem cells were seeded at 7,400 cells/ well in a 24-well plate, flat bottom format with expansion media, and incubated at 37°C, 5% CO₂ atmosphere until the cells attached to the bottom of the wells. The cells were then cultured for 7 days, with media changes every 2 days. The media conditions included expansion media and osteogenic media, both supplemented with either 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2, or 0 ng/mL BMP2. After 7 days of cell culture, the samples were subject to ALP Fluorometric assay kit (GeneTex) by following ab83371 Abcam protocol. Prior to subjecting cell lysates to the assay, the cultured cells had to be

prepared. Media was removed and the cell monolayers were washed twice in PBS before placing 200 μL /well of trypsin for 5 minutes at 37°C. DMEM with FBS was then added (300 μL /well) and the cell suspension was transferred into 1.5 mL Eppendorf tubes before being subject to centrifugation at 1500 rpm for 5 minutes using Eppendorf 5424R centrifuge. The cell pellet was resuspended in 100 μL complete DMEM and cell counting was performed using Countess 3 automated cell counter (Thermo Fisher). To fresh 1.5 mL Eppendorf tubes, 1×10^5 cells were added of each separate condition and centrifuged again using the same settings. The cell pellet was resuspended in PBS to ensure full removal of media and centrifuged once more. Subsequently, the generated cell pellet was resuspended in 100 μL assay buffer (provided in the ALP Fluorometric assay kit) and homogenised using 21 G needle syringe. The homogenate was centrifuged at 13,000 rpm for 3 minutes at 4°C. The supernatant was collected into a fresh 1.5 mL Eppendorf tube and kept on ice throughout the ALP assay protocol steps. Once the assay was concluded, fluorescence readings were taken using the Synergy Neo2 plate reader at 360 nm excitation and 440 nm emission.

5.2.6 REVERSE TRANSCRIPTION QUANTITATIVE POLYMERASE CHAIN REACTION (RT-QPCR)

Reverse transcription-quantitative PCR is a powerful technique for detecting gene expression levels by converting RNAs into cDNA, which can be fed to PCR amplification with a fluorescent probe for easy quantification¹⁶. As the profile of gene expression change over the course of hMSC differentiation, this technique can be used for tracking its timeline.

Human mesenchymal stem cells were seeded at 7,400 cells/well in a 24-well plate, flat bottom format with expansion media, and incubated at 37°C and 5% CO₂ atmosphere until the cells attached to the bottom of the wells. The cells were then cultured for 14 days, with media changes every 2 days. The media conditions used are as follows: expansion media and osteogenic media supplemented with either 25 ng/mL rhBMP2, 25

ng/mL *E.coli*-produced deGFP-BMP2, or completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2. After 14 days of cell culture, the samples were subject to RNA isolation and purification, a first step of RT-qPCR. PureLink RNA mini kit (Life technologies) was used for RNA extraction. Media was removed from the wells and cells were washed twice with PBS before applying 200 μ L/well of trypsin for 5 minutes at 37°C. DMEM with FBS was added (300 μ L/well) to inactivate trypsin and the cell suspension was transferred to 1.5 mL Eppendorf tubes. Centrifugation at 1500 rpm for 5 minutes in Eppendorf 5424R centrifuge was performed. The cell pellet was resuspended in 350 μ L RLT buffer (provided in PureLink RNA mini kit) with 2-mercaptoethanol and was subject to homogenization using a 21G needle syringe. The resulting homogenate was mixed with 70% ethanol and the RNA binding, washing, and elution steps were followed according to the manufacturer's instructions. Purified RNA was immediately used for cDNA preparation, ensuring the same concentration of starting RNA. High Capacity cDNA Reverse Transcription kit (Thermo Fisher) was used with the protocol provided by Applied Biosystems to carry out complementary DNA synthesis. The generated samples were stored in -80°C until further steps in qPCR were undertaken. The reactions for qPCR were all performed on ice to prevent sample degradation and TaqMan Gene Expression Assay kit (Thermo Fisher) was utilized to prepare the solutions. Reactions were set as 10 μ L volumes in MicroAmp Fast Optical 96-well plate (Applied Biosystems), containing a primer solution and a cDNA solution. Primer solution consisted of 5 μ L of TaqMan Master Mix and 0.5 μ L of TaqMan gene primer, whereas the cDNA solution contained 1 μ L of the cDNA template generated from the previous step, and 3.5 μ L of DEPC water. The plate was sealed with clear polypropylene adhesives for PCR plates (Starlab) and briefly centrifuged for 5 seconds at 1500 rpm using Eppendorf 5810 centrifuge with a plate adapter. Quantitative PCR was performed using Step One Plus Real-Time PCR system (Applied Biosystems) with the following steps: 2 minutes at 50°C, 10 minutes at 95°C, and 40 cycles of 15 seconds at

95°C and final 1 minute at 60°C. *GAPDH* was set as a housekeeping gene and the analysis was carried out using the double delta Ct method.

5.2.7 ALIZARIN RED S STAINING

Alizarin Red S (ARS) is an orange-red stain that can be used to visualize calcium deposits due to its formation of chelate complexes with calcium¹⁷. ARS can be used to distinguish between undifferentiated and osteoblast cells, due to the osteoblast hallmark of extracellular calcium deposits. ARS allows for qualitative assessment and once extracted, can be quantified spectrophotometrically.

Human mesenchymal stem cells were seeded at 7,400 cells/well in a 24-well plate, flat bottom format with expansion media, and incubated at 37°C and 5% CO₂ atmosphere until the cells attached to the bottom of the wells. The cells were then cultured for 28 days, with media changes every 2 days. The media conditions used are as follows: expansion media and osteogenic media supplemented with either 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, or completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2. After 28 days of cell culture, the ARS staining was performed. Powdered ARS (1.369g) (Sigma Aldrich) was dissolved in ~75 mL of distilled water and pH adjusted to 4.2. The volume was then corrected to 100 mL with distilled water. The cells were fixed with >98% ice cold methanol for 10 minutes once cell media was removed. Then, a wash step with PBS was performed. ARS at 0.5 mL/ well was added on top of the cell monolayers and incubated, protected in aluminum foil from light, for 20 minutes at room temperature with gentle rocking. The stain was removed, and four PBS washes were carried out. Images using Leica S9i digital stereo microscope were taken. Afterwards, the ARS was removed with 500 µL/well of 10% w/v cetylpyridinium chloride (Sigma Aldrich) for 1 hour with gentle rocking. The solutions were then transferred (200 µL) into 96-well plate with clear bottom for absorbance quantification using Synergy Neo2 plate reader (BioTek).

5.2.8 STATISTICAL ANALYSIS

Cell culture data in this chapter were obtained from at least three patient numbers, each with three technical repeats in any given experiment. One-way Analysis of Variance (ANOVA) or the two-tailed, unpaired student t-test (qPCR data) were applied in GraphPad Prism 9 and a value of (*) $p < 0.05$ was considered statistically significant. (**) denotes $p < 0.01$, (***) denotes $p < 0.001$ and (****) denotes $p < 0.0001$.

5.3 RESULTS AND DISCUSSION

5.3.1 VISUALISING DEGFP-BMP2 CHIMERA

It was important to observe if the chimera interacted with hMSCs and determine if any internalisation occurred. The inherent fluorescence properties of deGFP-BMP2 were utilised to visualise interactions between deGFP-BMP2 and hMSCs using confocal microscopy. Additionally, the cytoplasm of hMSCs was stained with CellMask Orange to provide information on co-localisation on cells. Human mesenchymal stem cells with 400 $\mu\text{g}/\text{mL}$ deGFP-BMP2 in media were imaged after 20, 80 and 120 minutes of incubation (Figure 5.2). Although the fluorescence from deGFP-BMP2 appears dim, small particle-like features could be seen at the plasma membrane of hMSCs, (green channel of first column in Figure 5.2). After 20 minutes of incubation between the chimera and hMSCs, the fluorescent speckles corresponding to deGFP-BMP2 appear to be scattered throughout the imaging field of view. However, after 80 minutes of incubation, the highest density of the green fluorescent speckles show to be mostly located intracellularly, when it is compared to the CellMask Orange channel (second column in Figure 5.2). This likely indicates that the chimera is first internalised by the cells between 20 and 80 minutes of incubation. It has been shown that BMP2 is taken up from the plasma membrane along with its receptors *via* caveolae and clathrin-mediated endocytosis, and together are sorted at late endosomes¹⁸. The interaction imaged at 120 minutes, shows that the amount of chimera speckles is reduced but their appearance is larger and brighter. The speckles also appear to be closer to the centre of the cell. It can be speculated that at this time the internalised chimera could have been condensed at the multivesicular body, which acts as an intermediate in the degradation pathway¹⁹. Alternatively, BMP2 could still be associated with the receptors, undergoing slow recycling; a process of budding and fusion involving recycling endosomes located at the perinuclear region²⁰. However, to confirm this, specific organelles would require additional staining to perform correlations with

deGFP-BMP2. Flow cytometry was another method that was applied to understand the interactions between the chimera and hMSCs (Figure 5.3 and 5.4).

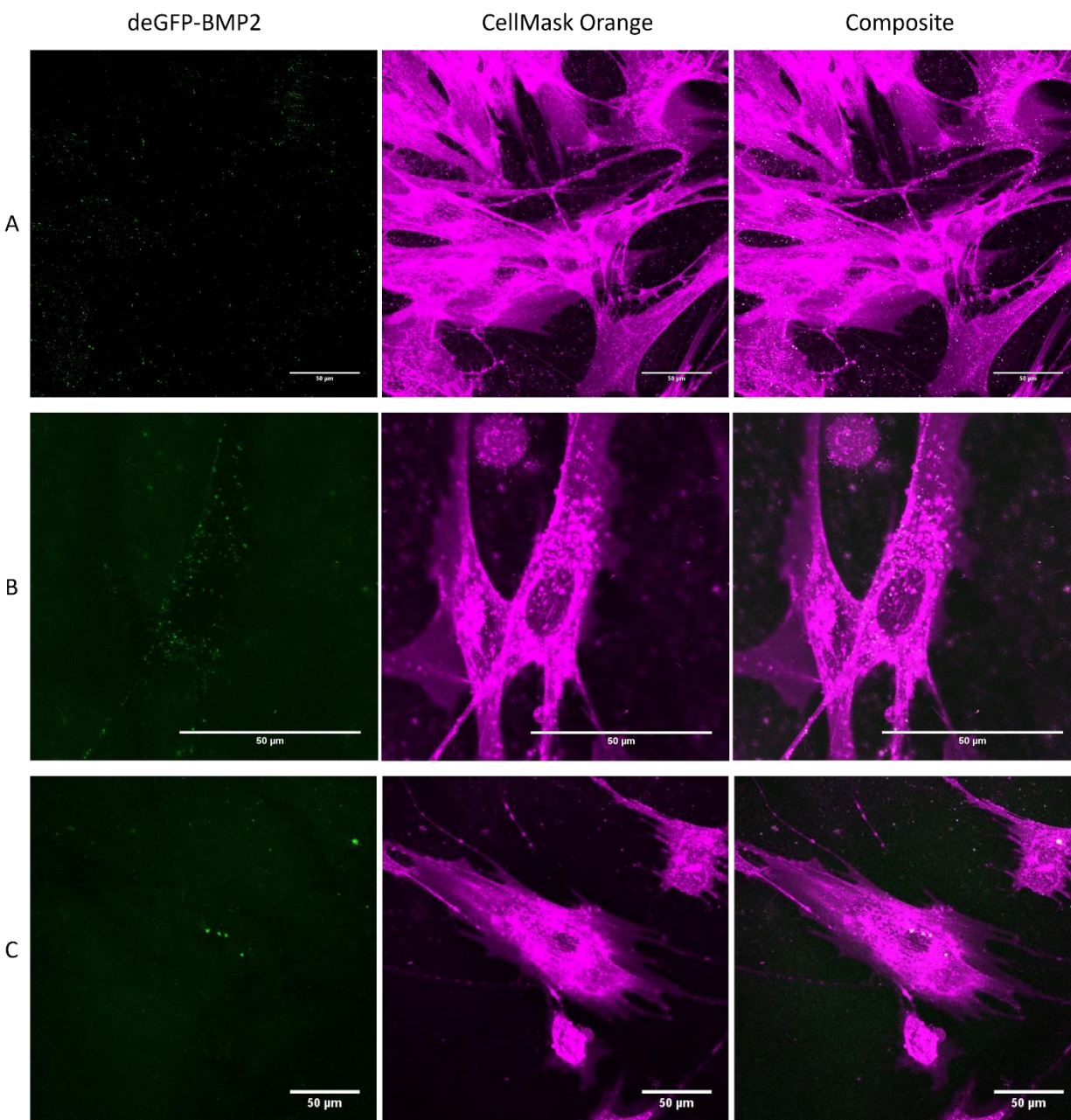


Figure 5.2 Z-stack confocal images of hMSCs incubated with 400 µg/mL *E.coli*-produced deGFP-BMP2 chimera for three different time periods of A: 20 minutes, B: 80 minutes and C:120 minutes. First column images deGFP-BMP2 fluorescence (shown in green), second column corresponds to CellMask Orange stain (shown in magenta) and the third column is the composite of the two (chimera shown in white for higher contrast). Scale bar: 50 µm

Untreated hMSCs from each tested patient were utilised for the gating strategy by firstly selecting a population of cells (Figure 5.3 column A). The use of a density map on forward scatter (FSC-A) vs side scatter (SSC-A) enabled the identification of most dense population (regions of red). Although the most dense regions in column A are present below 50K of FSC-A and SSC-A, the individual particles of those parameters relate mostly to debris, and therefore, are discarded from gate. The population of hMSCs in patient 1 (P1) is at a lower FSC-A in comparison to patient 2 (P2) data, whereas both of these populations are present in patient 3 (P3). The lower FSC-A demonstrate smaller cell size and there might be several reasons why the two populations exist. The size of hMSCs is influenced by factors such as the surface they grow on, their cell cycle phase and their passage number²¹. Although hMSCs past their 6-passage number were not used here in order to maintain stemness²², the discrepancy in passage numbers between patient numbers could bring about cell size variation²³. Another reason why two populations are observed in column A of P3 in Figure 5.3, might be due to cell death. As cells experience apoptosis or necrosis, their morphology changes drastically, including the loss of the overall cell shape²⁴. Here, the examined cells were suspended in PBS, rather than DMEM to minimise background fluorescence. Since cells cannot be suspended in PBS for an extended period of time, cell death could be arising at the time of taking these measurements. Stains such as propidium iodide would have been beneficial in determine the viability of cells. Further gating was performed on both hMSCs populations, to determine if any fluorescence differences exist between them. Doublet discrimination was carried out on FSC-A vs FSC-H plot on the selected hMSCs populations (Figure 5.3, column B), followed by selection of GFP positive and negative cells using FITC-A channel (Figure 5.3, column C). The difference in GFP selection of both populations in P3 was minimal. Once gating was performed, the resulting GFP data has been presented as histograms of FITC-A GFP vs cell count, with untreated hMSCs in cyan, and three

technical repeats of each condition in pink, violet, and yellow-green (Figure 5.4). The percentages of GFP positive cells are stated in Table 5.2.

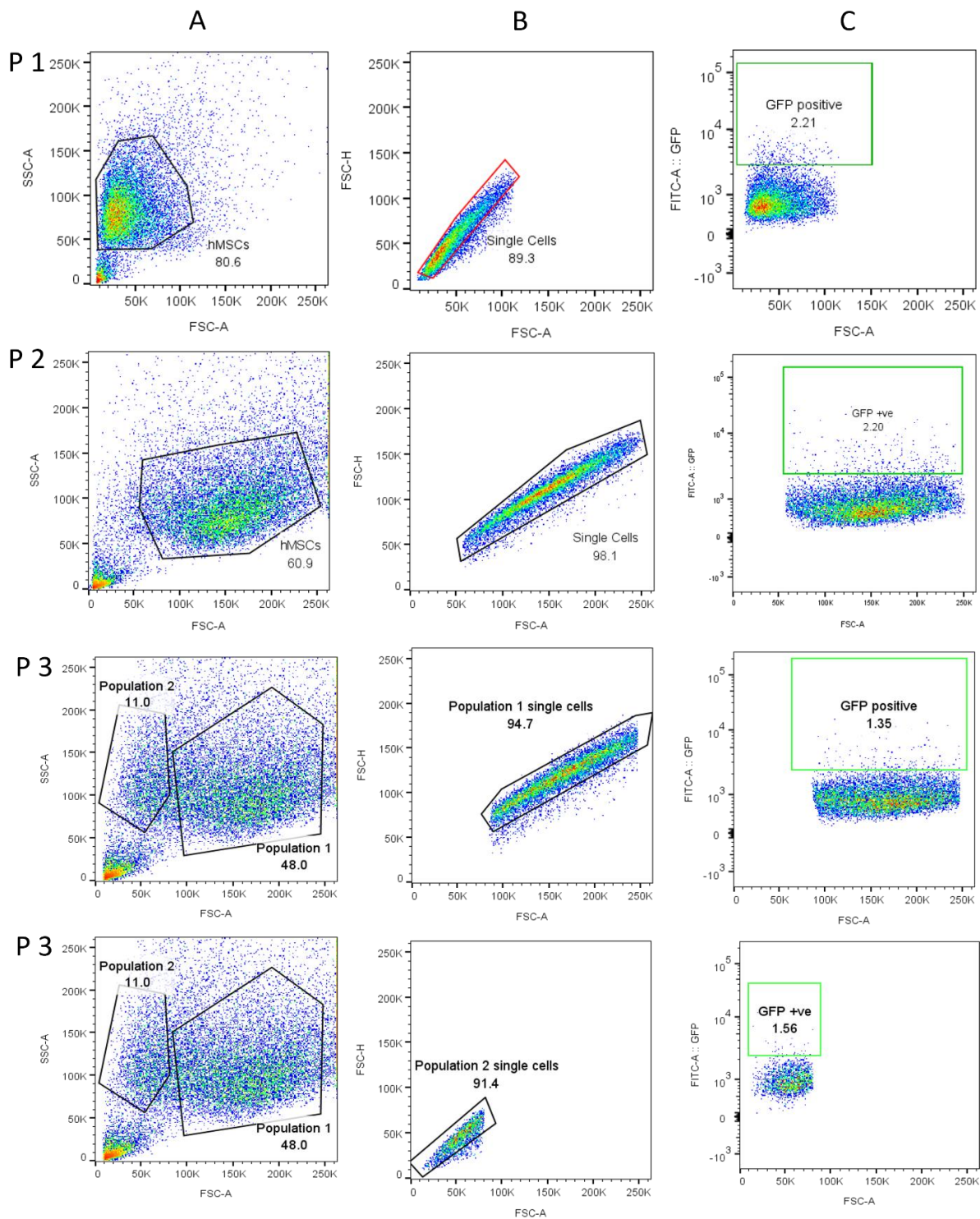


Figure 5.3 Flow cytometry gating strategy displaying density plots of 10,000 events (low density blue to high density red). Flow cytometry experiment was performed with three patient numbers (P1-P3) each with three technical repeats. Non-labelled hMSCs were utilised to designate the following A: populations using FSC-A vs SSC-A parameters, B: single cells using FSC-A vs FSC-H and C: GFP fluorescence using FSC-A vs FITC-A.

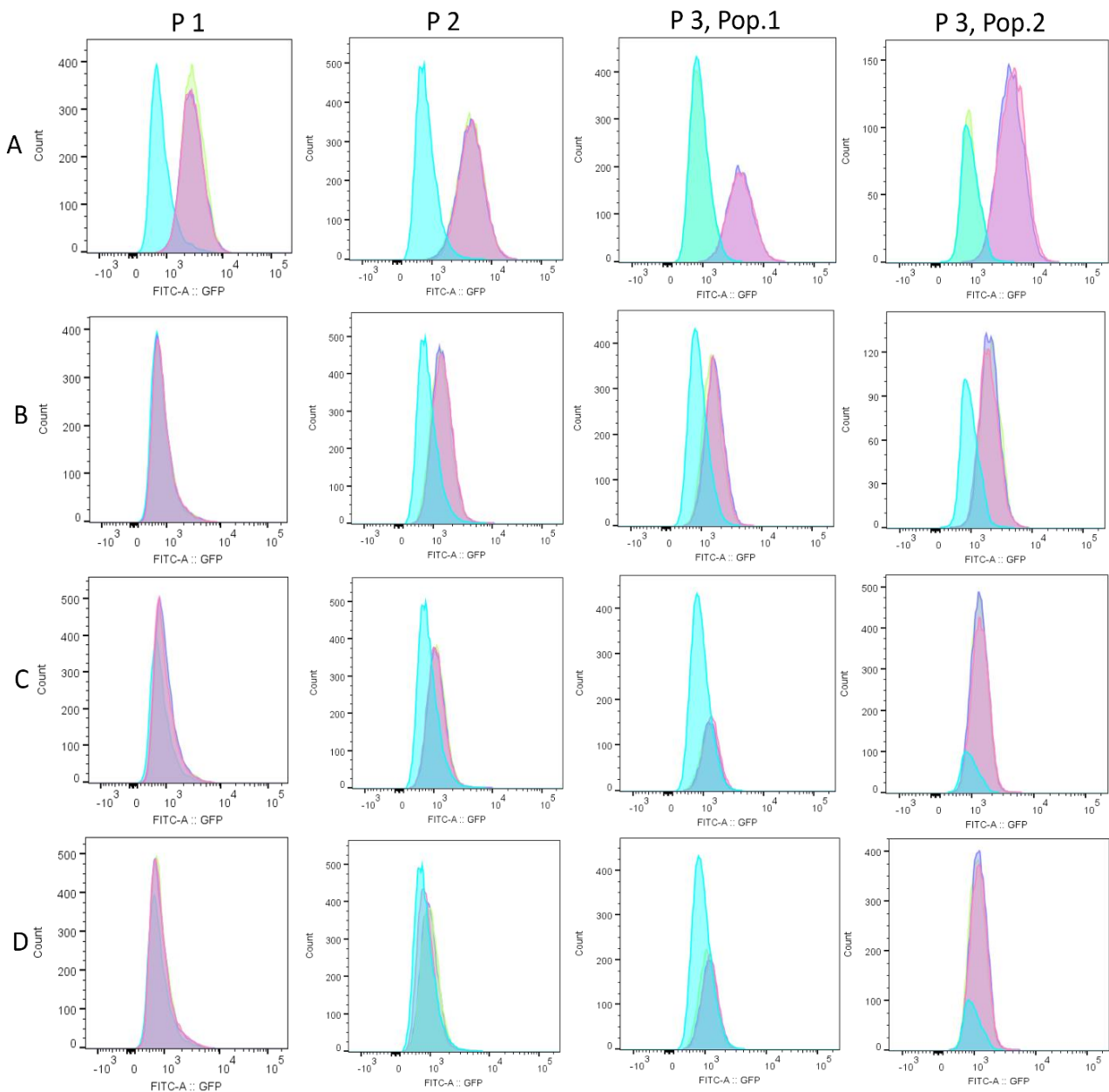


Figure 5.4 Flow cytometry histograms of GFP fluorescence presented as FITC-A GFP vs cell count for each tested condition. Histograms have been constructed after population and single cell gating. Flow cytometry experiments were performed with three patient numbers (P1-P3) and both detected population of cells in patient 3 are included (Pop.1 and Pop.2). A: 0.1 mg/mL *E.coli*-produced deGFP-BMP2 chimera. B: 0.1 mg/mL EGFP. C: 3.54 x 10⁻³ mg/mL deGFP-BMP2 produced by CFES. D: 3.54 x 10⁻³ mg/mL deGFP produced by CFES. Unlabelled hMSCs are presented in cyan, whereas three technical repeats of each condition are presented as pink, violet, and yellow-green.

	Patient 1					Patient 2				
	Repeats (%)			Average	Std.dev	Repeats (%)			Average	Std.dev
deGFP-BMP2	46.60	46.90	50.40	47.97	1.72	91.60	89.70	89.10	90.13	1.07
EGFP	2.69	2.45	2.38	2.51	0.13	9.46	9.83	7.64	8.98	0.96
CFES deGFP-BMP2	2.50	2.53	3.04	2.69	0.25	3.34	3.14	3.30	3.26	0.09
CFES deGFP	2.49	2.67	2.68	2.61	0.09	2.41	2.22	1.96	2.20	0.18

	Patient 3, Population 1					Patient 3, Population 2				
	Repeats (%)			Average	Std.dev	Repeats (%)			Average	Std.dev
deGFP-BMP2	87.1	87.9	1.14	58.71	40.71	90.4	89.2	1.63	60.41	41.57
EGFP	15.1	12.9	9.83	12.61	2.16	37	32.6	29.1	32.90	3.23
CFES deGFP-BMP2	2.52	2.78	4.17	3.16	0.72	2.78	3.36	4.28	3.47	0.62
CFES deGFP	1.26	1.73	1.96	1.65	0.29	1.45	1.62	1.98	1.68	0.22

Table 5.2 Summary of flow cytometry results of GFP positive populations, expressed as percentages for each repeat per condition. The averages of the repeats and standard deviation is also noted.

In Figure 5.4 and Table 5.2, the largest shift in fluorescence intensity can be noted between untreated (cyan) and hMSCs treated with 0.1 mg/mL *E.coli*-produced deGFP-BMP2 chimera in P1 and P2 (row A). This shift is also seen in P3, but only in two technical repeats, with 87.1%, and 87.9% of the cells being GFP positive, in comparison to the 1.14% in the one technical repeat in population 1 (90.4%, 89.2% and 1.63% in population 2). The histogram of the third technical repeat lines up with the untreated sample signal with the average GFP positive cell count of $(1.83 \pm 0.38)\%$, which could indicate a possible inaccuracy during treatment dispensing, however, to confirm this, additional sample numbers would have to be measured. Media incubation with EGFP has yielded a shift in fluorescence in P2 and P3 hMSCs (row B), although it is much lower than deGFP-BMP2 signal. There are no specific EGFP receptors on hMSCs surface, therefore, the interactions between EGFP and hMSCs are expected to be minimal. The observed fluorescence shift is possibly the result of non-specific EGFP-membrane interactions and consequent internalisation. Populations 1 and 2 from P3 do not appear to be significantly different in their GFP positive percentages, except in the EGFP signal, with $\sim 20.3\%$ difference. The signal generated from CFES-produced deGFP-BMP2 (row C) is significantly reduced in comparison to *E.coli*-produced deGFP-BMP2 or even EGFP, however, in P2 and P3 (3.26%, 3.16% and 3.47% respectively) it does appear elevated than its negative control

counterpart, CFES-produced deGFP (2.20%, 1.65% and 1.68%) (row D). Although this difference is minimal and therefore further optimisation is necessary to quantify this result. The signal difference between CFES-produced chimera and *E.coli*-produced chimera could be attributed to the initial media supplementation concentration, which was much lower for CFES-made protein. This was due to the difficulty in scaling-up cell-free protein synthesis concentration for the use in the flow cytometry experiments. It would have been easier to compare the results if *E.coli*-produced chimera concentration was reduced to the same media supplementation levels as CFES-produced chimera. Generally, the higher fluorescence signal from *E.coli*-produced deGFP-BMP2 and CFES-produced deGFP-BMP2, in comparison to their negative control counterparts signify an ongoing interaction between the engineered BMP2 protein with hMSCs.

5.3.2 DOWNSTREAM SIGNALLING OF DEGFP-BMP2

In light of the visualised interaction between the chimera and hMSCs, it was important to determine whether BMP2 was capable of binding to the cell receptors BMPRI and BMPRII, and therefore, whether it was biologically active. Upon binding of BMP2 to the receptors to form heterotetrametric complexes, the stimulated serine kinase activity of the receptor induces phosphorylation of the receptor-regulated Smad proteins (R-Smads) including Smad1, Smad5 and Smad9²⁵. Once the Smad proteins are phosphorylated, they form a complex with Smad4 and translocate to the nucleus to induce upregulation of osteogenic factors. Accordingly, the BMP2 induced phosphorylation of Smad was tested on a semi-quantitative Western blot (Figure 5.5 and Table 5.3).

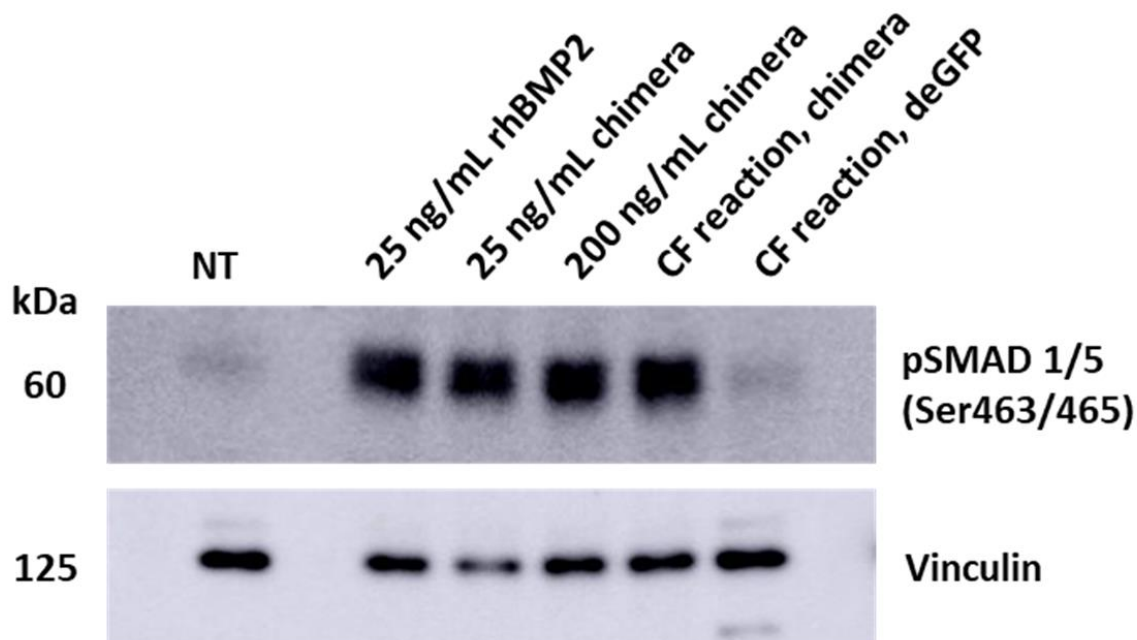


Figure 5.5 Western blot assessing the phosphorylation of Smad proteins 1 and 5 in cell lysates. hMSCs were incubated for 4 hours with 25 ng/mL rhBMP2 (positive control), 25 ng/mL and 200 ng/mL *E.coli*-produced deGFP-BMP2 chimera, and complete 10 μ L CFES reactions of deGFP-BMP2 chimera and deGFP per 1 mL of media. Non-treated hMSCs (NT) as negative control. Vinculin (125 kDa) was used as a loading control.

Lane	pSMAD (Raw)	Vinculin (Raw)	Norm. factor	Norm. signal
NT	160	2604	0.92	174.44
25 ng/mL rhBMP2	3197	2258	0.80	4019.61
25 ng/mL chimera	3637	1748	0.62	5907.00
200 ng/mL chimera	3277	2839	1.00	3277.00
chimera CF reaction	2787	2627	0.93	3011.91
deGFP CF reaction	408	2768	0.97	418.47

Table 5.3 Quantification of the Western blot using GelAnalyzer software. Lane normalisation was calculated from observed vinculin raw signal and the highest observed vinculin raw signal. The resulting normalisation factor was then used with the pSmad raw values to generate a normalised signal. GelAnalyzer 19.1 (www.gelanalyzer.com) by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc.

Both *E.coli*-produced and CFES-produced deGFP-BMP2 induced phosphorylation of the Smad1 and Smad5, as evidenced in Figure 5.5. This demonstrates the physiological activity of deGFP-BMP2 chimera and its ability to effectively bind to BMP2 receptors.

The highest level of phosphorylated Smad were obtained using 25ng/mL of *E.coli*-produced chimera, which was found to have a more profound effect in comparison to the positive control of commercial recombinant human BMP2 (rhBMP2). An increase in deGFP-BMP2 from 25 ng/mL to 200 ng/mL in media did not generate a higher amount of phosphorylated Smad, suggesting that BMP2 concentration as low as 25 ng/mL is sufficient to trigger downstream signalling. As a negative control, CFES-produced deGFP showed minimal effect on Smad phosphorylation, as the detected levels were more comparable with the levels from untreated hMSCs.

5.3.3 DIFFERENTIATION ASSAYS

There are numerous assays to determine the differentiation lineage commitment of hMSCs, with alkaline phosphatase (*ALP*) being a common osteogenesis marker. In this assay 4-methylumbelliferyl phosphate disodium salt (MUP) is added as a substrate to produce a fluorometric outcome. The strength of the signal depends on the amount of ALP present in the sample. Here, ALP assay was carried out after 7 days of hMSCs culture in expansion and osteogenic media with different versions of BMP2 (Figure 5.6). It is evident that ALP activity is significantly elevated for the osteogenic media conditions with all the versions of BMP2, including a positive control rhBMP2 (light grey), *E.coli*-produced deGFP-BMP2 and CFES reaction containing the produced deGFP-BMP2 (light blue and light red respectively). However, it is also apparent that ALP activity can be increased when hMSCs are cultured in DMEM supplemented only with dexamethasone, ascorbic acid and β -glycerophosphate, without the addition of BMP2 (pink). These four conditions have heightened ALP activity to similar extent. On the contrary, significant ALP activity is detected from samples where DMEM was supplemented with *E.coli*-produced and CFES-produced chimera only (dark blue and dark red respectively). Therefore, it is possible to trigger osteogenesis with the in-house-made chimera as the only supplement in the culture media. Interestingly, no significant amount of ALP can be detected for hMSCs cultured in DMEM with rhBMP2 only (dark grey).

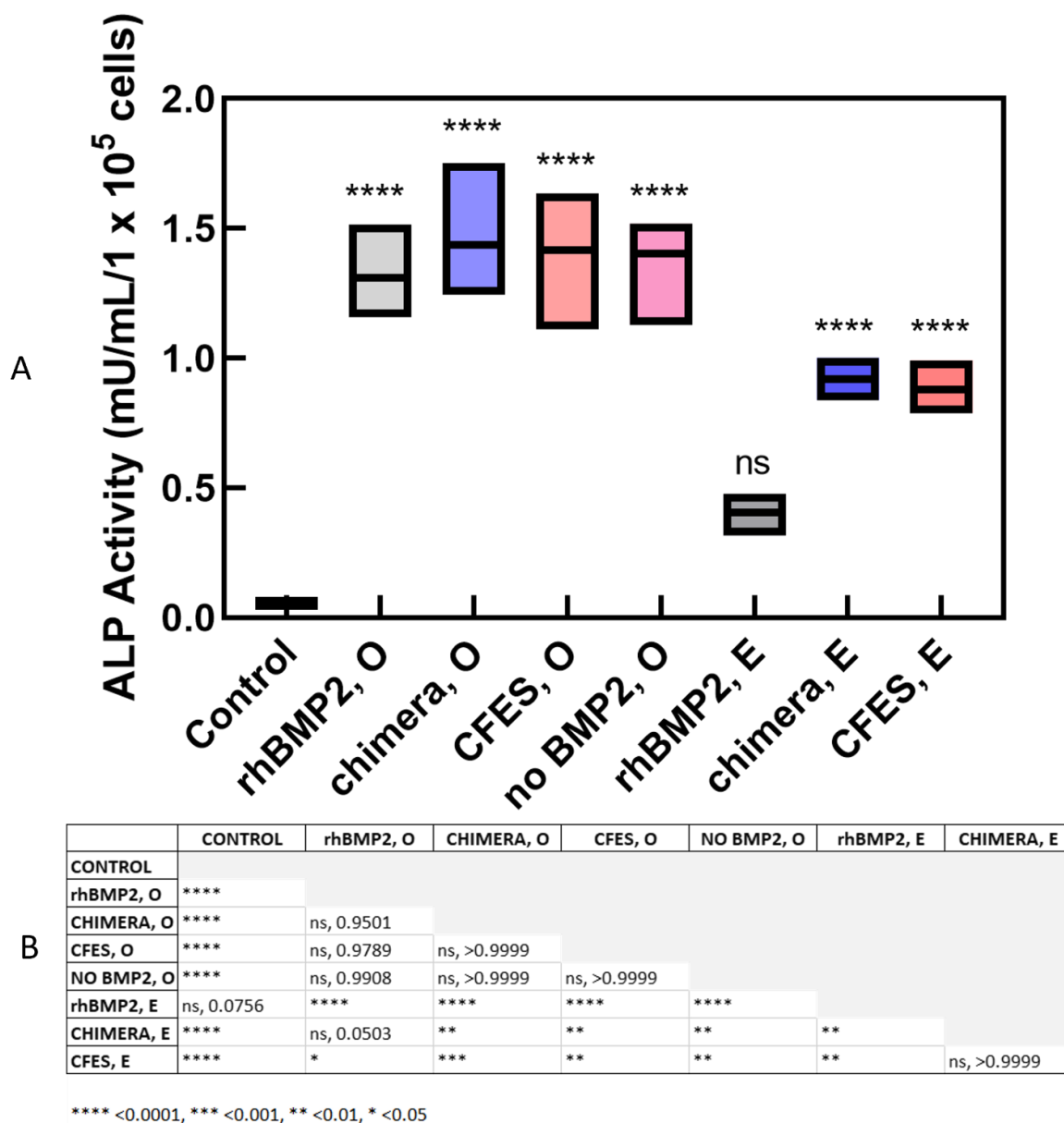


Figure 5.6 ALP activity measured from lysates originating from hMSCs incubated for 7 days in either expansion media denoted as ‘E’ or osteogenic media denoted as ‘O’, supplemented with 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2, or 0 ng/mL BMP2. ALP activity unit is defined as the amount of enzyme causing the hydrolysis of 1 μ mol of MUP per minute at pH 10 and 25°C. A: Minimum and maximum floating bar graph of ALP activity with mean values at the middle line. Data accumulated from three technical repeats of three patient samples. Number of asterisks correspond to the degree of statistical significance. B: P-values calculated from ANOVA test.

Alongside ALP, osteogenesis progression was also examined using quantitative PCR after 14 days of cell culture. Here, mRNA levels of four different genes, upregulated as a response to osteogenic differentiation, were investigated (Figure 5.7). The expression levels of these genes were normalized to the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*), a regulatory enzyme in glycolysis²⁶. The resulting data was then normalized to the untreated cells, cultured for 14 days in expansion media. This method of analysis is referred to as $\Delta\Delta C_t$.

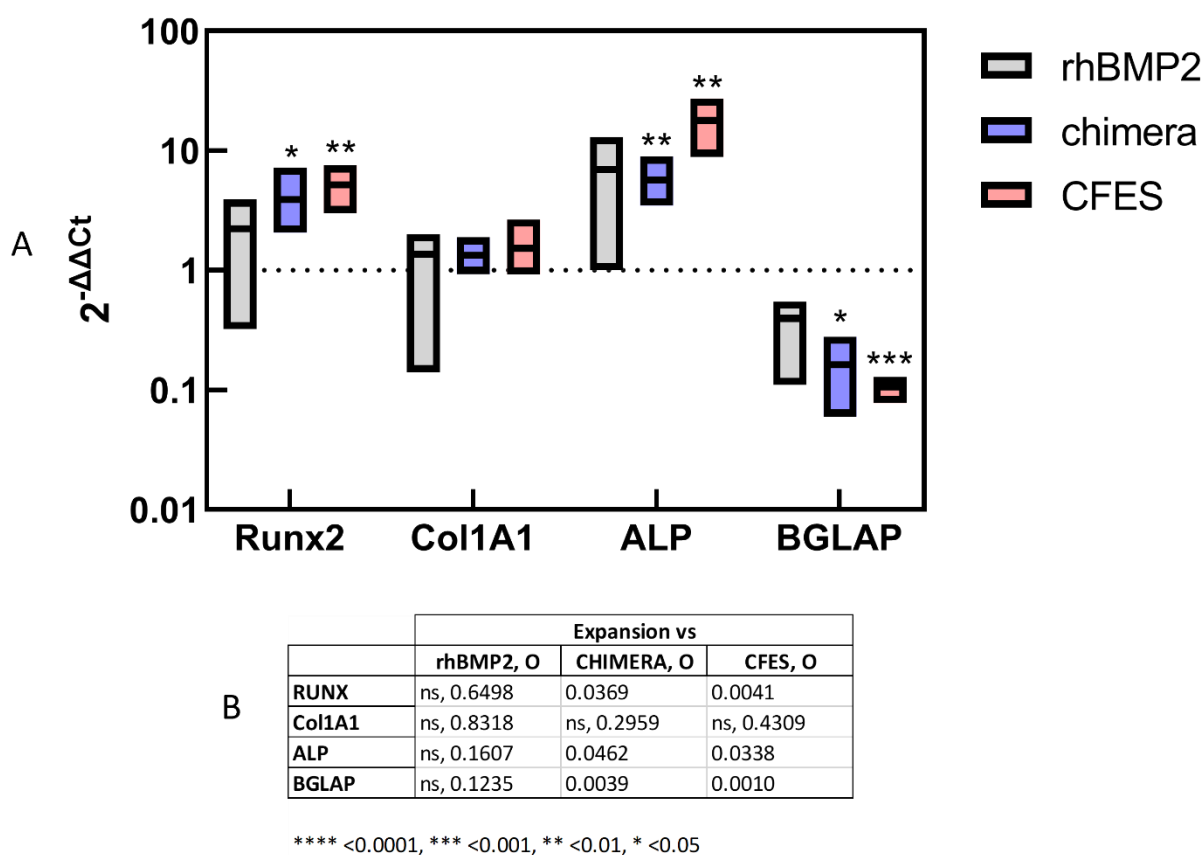


Figure 5.7 Quantitative PCR of *Runx2*, *Col1A1*, *ALP* and *BGLAP* from hMSCs conditioned for 14 days in osteogenic media containing 25 ng/mL rhBMP2, 25 ng/mL *E.coli*-produced deGFP-BMP2, completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2. The results are relative to a housekeeping gene *GAPDH* in undifferentiated hMSCs cultured with expansion media. The data is expressed as expression fold change ($2^{-\Delta\Delta C_t}$) following double delta Ct analysis. A: Minimum and maximum floating bar graph with mean values denoted by the middle line. Data accumulated from three technical repeats of three patient samples. B: P-values calculated from two-tailed, unpaired student t-test, based on $\Delta\Delta C_t$ values.

In Figure 5.7, the calculated fold change means (middle line of the floating bars) from cells cultured with rhBMP2 and both *E.coli*- and CFES-made chimera are higher in *RUNX2*, *Col1A1* and *ALP* genes. This suggests the three versions of BMP2 brought about upregulation of these three genes. Although, no statistical difference in the fold change between expansion samples and rhBMP2 cultured cells was calculated. The lack of statistical difference between expansion and osteogenic conditions is also seen with the *Col1A1* gene, even when the fold change means are above 1. The largest fold change from expansion samples can be observed from chimera generated by a CFES reaction, where a statistical difference in *RUNX2*, *ALP* and *BGLAP* genes is seen. This indicates the most prominent osteogenic potential. *BGLAP* gene appears to be downregulated in hMSCs treated with rhBMP2, *E.coli*- or CFES-made chimera as the fold change data is below 1. Osteocalcin (*BGLAP*) is a late osteogenic marker which is upregulated during bone mineralisation and matrix synthesis by mature osteoblasts²⁷. As qPCR was conducted 14 days after osteogenesis was induced, it would be unlikely for *BGLAP* to be upregulated at this time-point. Conversely, *ALP* and *Col1A1* are early osteogenesis markers, firstly observed during osteoprogenitor to pre-osteoblast cell phase²⁸. Runt-related transcription factor 2 (*RUNX2*) is a master regulator that leads upregulation of other markers and is expressed throughout the whole process of osteogenic differentiation²⁹. The gene regulation profile in Figure 5.7 resembles that of early osteogenesis.

During bone formation, the process of calcification takes place at nucleation sites described as matrix vesicles that build up calcium and inorganic phosphate³⁰. Over time, hydroxyapatite ($\text{Ca}_{10}(\text{PO}_4)_6(\text{OH}_2)_2$), a major component of bone, begins to form at these nucleation sites. Alizarin Red S stain enables to identify these calcium-rich deposits when cells undergo osteogenic differentiation (Figure 5.8). A positive reaction where Alizarin Red S forms chelate complexes with calcium resulting in a bright red stain can be observed in hMSCs cultured in osteogenic media containing rhBMP2, *E.coli*-produced deGFP-BMP2 and CFES-produced deGFP-BMP2 (Figure 5.8 B,C,D). Images of samples

containing expansion media alone (A) and CFES-produced deGFP-BMP2 in expansion media (E) show cells with a pink hue, indicating lack of calcium deposits. This could suggest that CFES-produced deGFP-BMP2 alone is not sufficient for calcium deposition by cells. The images correlate to the absorbance spectra obtained post cetylpyridinium chloride extraction (F), where the absorbance peak at 550 nm is of importance. The 550 nm peak of CFES-produced deGFP-BMP2 in osteogenic media is the highest absorbing out of the tested conditions. The amount of ARS present in each sample was then quantified by plotting a standard curve with known ARS concentrations. The results containing three technical repeats from three separate patients are summarised in Figure 5.9.

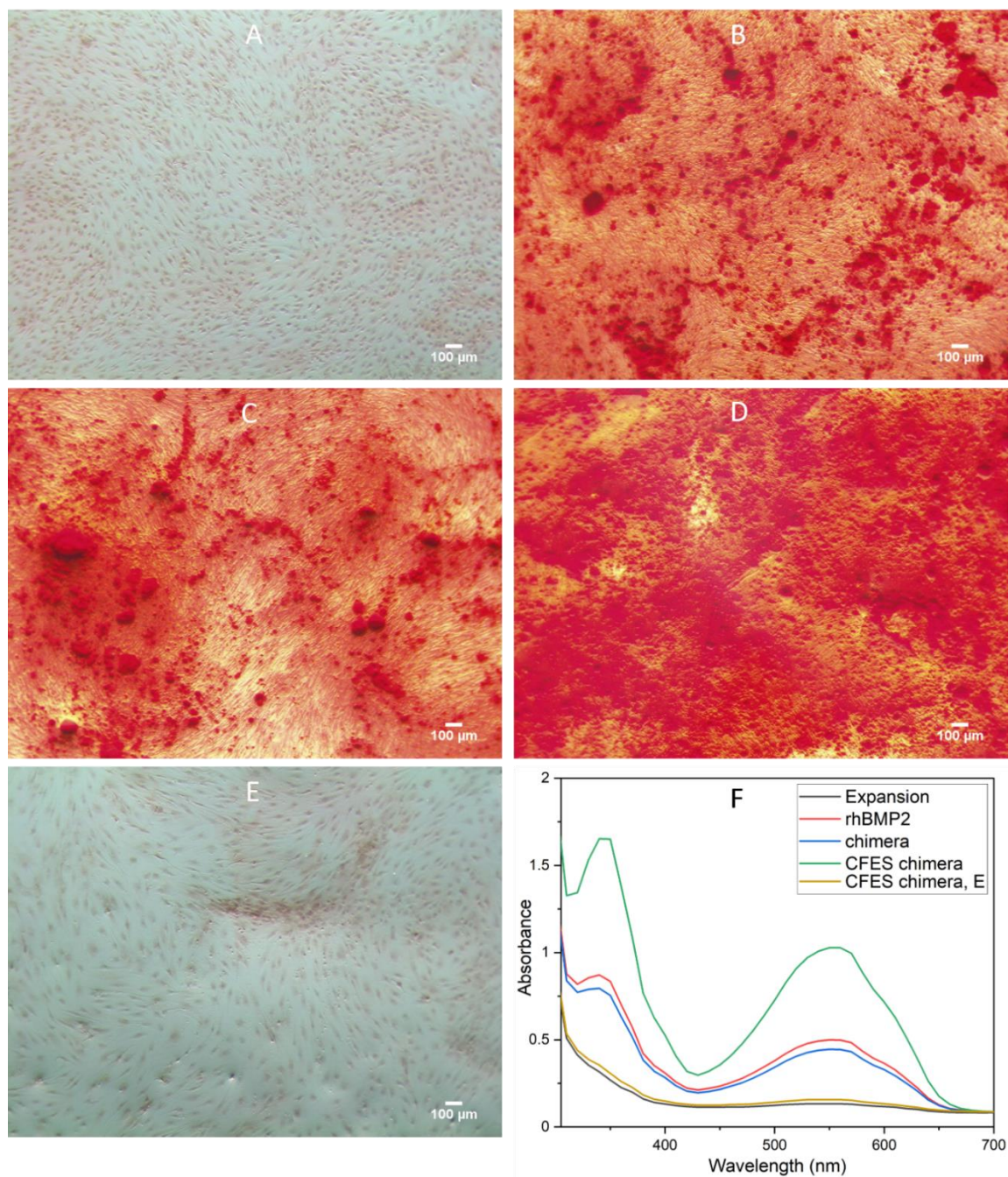


Figure 5.8 Representative Alizarin Red S stain images (A-E) and absorbance scan (F) from hMSCs cultured for 28 days in osteogenic differentiation media. A: expansion media. B: osteogenic media with 25 ng/mL rhBMP2. C: osteogenic media with 25 ng/mL *E.coli*-produced deGFP-BMP2. D: completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2 with osteogenic media. E: completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2 with expansion media. Positive pigment formation in the presence of calcium deposits appears as orange to red in colour. Negative staining is visualized as pink. Scale bar at 100 µm. F: Absorbance spectral scan after ARS was dissolved from each condition. Expansion media denoted as 'E'.

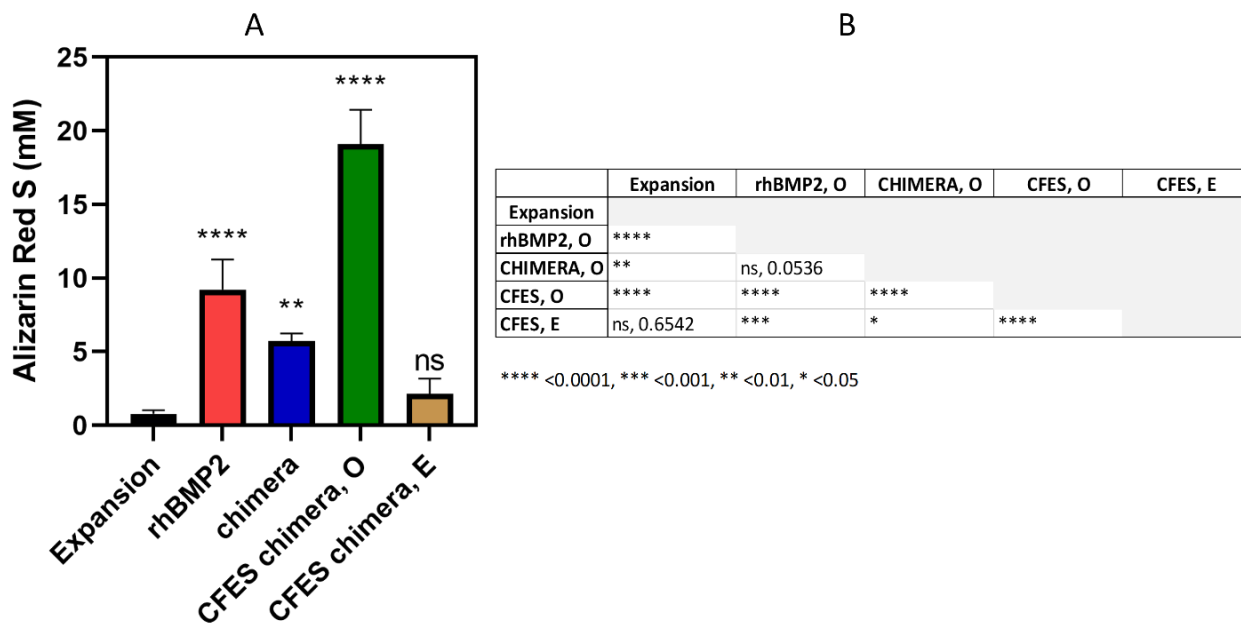


Figure 5.9 Alizarin Red S quantified from each tested condition after the stain was extracted and standard curve was performed. A: bar graph of ARS amount in mM from hMSC cultured in expansion (black), osteogenic media with 25 ng/mL rhBMP2 (red), osteogenic media with 25 ng/mL *E.coli*-produced deGFP-BMP2 (blue), completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2 with osteogenic media (green), and completed CFES reaction with the final estimated concentration of 25 ng/mL of produced deGFP-BMP2 with expansion media (yellow). Data accumulated from three technical repeats of three patient samples. B: P-values calculated from ANOVA test

The quantification of ARS in Figure 5.9 verifies the observations made from ARS staining images. ARS concentration from rhBMP2, *E.coli*- and CFES-produced deGFP-BMP2 in osteogenic media show to be statistically significant in comparison to the negative control of expansion media only. The (2.14 ± 0.85) mM ARS from CFES-produced deGFP-BMP2 in expansion media failed to generate a statistically significant result. The highest amount of ARS obtained at (19.10 ± 1.90) mM was generated by CFES-produced chimera in osteogenic media. This was higher than the positive control of rhBMP2, which could suggest that the cell extract supplemented with numerous ions and additives could be contributing to the enhanced calcium deposition, although, it alone is not sufficient to produce this effect as seen by the yellow bar.

5.4 CONCLUSIONS AND FURTHER WORK

In this chapter the interactions of deGFP-BMP2 with human mesenchymal stem cells along with the effects of such interactions, were investigated. Firstly, by employing confocal microscopy, the fluorescence of the chimera was visualised on membranes of hMSCs, followed by endocytosis over time. The exact timing of internalisation could be determined by time-lapse imaging. It could also be beneficial to use rhBMP2 labelled with a dye to compare the interactions alongside the chimera, and implement a negative control of EGFP in future studies. By labelling early and recycling endosomes, co-localisation of deGFP-BMP2 with organelles could be carried out once internalised. Flow cytometry revealed that hMSCs incubated with *E.coli*-made deGFP-BMP2 increased the fluorescence of the cells over the EGFP controls. Higher fluorescence, but to a lesser degree, was also noted in hMSCs labelled with CFES-made chimera in comparison to CFES-made deGFP. The difference in fluorescence labelling between *E.coli* made and CFES-made chimera can be attributed to differences in concentrations. Two populations of hMSCs were observed in FSC-A vs SSC-A, and although no significant divergence in fluorescence was noted between them, it is likely that ‘population 2’ represented dead cells. Viability markers, such as propidium iodide, would enable distinction between live and dead cells. When trypsin was added to hMSCs to aid detachment from well-plates for transfer into FACS tubes, the cells were resuspended in PBS. To maintain health of the cells, low-fluorescence media should have been used instead. Flow cytometry could also be used to monitor osteogenesis progression by using fluorescent antibodies against specific surface markers such as CD73 and CD90 on undifferentiated hMSCs³¹, and CD10, which is increasingly expressed by osteogenically differentiated hMSCs³². Western blotting of phosphorylated Smad proteins confirmed the ability of the chimera to bind to the receptors of BMP2 on hMSCs surface (BMPRI and BMPRII) and trigger downstream signalling which leads to osteogenic differentiation. This effect was most noticeable with the *E.coli*-produced chimera, surpassing that of a positive control. ALP assay also

supported the observed osteogenic differentiation triggered by deGFP-BMP2, with a performance that was commensurate with commercial rhBMP2. ALP activity was also significantly higher than expansion control, from cells incubated in expansion media supplemented with deGFP-BMP2 alone (both *E.coli*- and CFES-made), suggesting this construct is capable of upregulating ALP without any other osteogenic agents. Additionally, qPCR showed regulation of several key genes involved in osteogenesis. Here, the calculated means of fold change after double normalisation of the early marker of osteogenesis, *ALP*, and a stable marker, *RUNX2*, were higher, suggesting upregulation of those genes after cell culture with *E.coli* or CFES- made deGFP-BMP2. A late osteogenic marker, *BGLAP* was downregulated, which coordinates with the timing of the qPCR test. For further work, the number of patient samples should be increased and the comparison should be conducted against hMSCs treated in osteogenic media without BMP2. Although *GAPDH* is the most commonly used reference gene in qPCR, it has been shown that it can become unstable in certain studies due to it being a metabolic protein³³. The best approach would be to generate a panel of housekeeping genes and determine the most stable reference, specific to own study. Successful calcium deposition by hMSCs cultured with the chimera in osteogenic media was shown, however, CFES-made deGFP-BMP2 in expansion media alone was insufficient in producing significant amount of extracted Alizarin Red. In future study, additional controls could be implemented, such as rhBMP2 and chimera in expansion media. The deposits that were detected in some of the samples could be further analysed for presence of hydroxyapatite crystals by x-ray powder diffraction. In summary, deGFP-BMP2 produced by either purifying from *E. coli* or through production from the CFES, drove osteogenic differentiation of hMSCs with a potency equivalent or even beyond that of the commercial rhBMP2.

5.5 BIBLIOGRAPHY

1. Ryoo, H. M., Lee, M. H. & Kim, Y. J. Critical molecular switches involved in BMP-2-induced osteogenic differentiation of mesenchymal cells. *Gene* 366, 51–57 (2006).
2. Kang, Q. *et al.* Characterization of the distinct orthotopic bone-forming activity of 14 BMPs using recombinant adenovirus-mediated gene delivery. *Gene Therapy* 2004 11:1711, 1312–1320 (2004).
3. Lin, S., Kh Svoboda, K., Feng, J. Q. & Jiang, X. The biological function of type I receptors of bone morphogenetic protein in bone. *Bone Research* 4, (2016).
4. Koenig, B. B. *et al.* Characterization and Cloning of a Receptor for BMP-2 and BMP-4 from NIH 3T3 Cells. *Molecular and Cellular Biology* 14, 5961–5974 (1994).
5. Wu, M., Chen, G. & Li, Y. P. TGF- β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. *Bone Research* 4, 1–21 (2016).
6. Bonor, J. *et al.* Initiation of BMP2 Signaling in Domains on the Plasma Membrane. *Journal of Cell Physiology* 227, 2880–2888 (2012).
7. Otto, F. *et al.* Cbfa1, a Candidate Gene for Cleidocranial Dysplasia Syndrome, Is Essential for Osteoblast Differentiation and Bone Development. *Cell* 89, 765–771 (1997).
8. Ponzetti, M. & Rucci, N. Osteoblast Differentiation and Signaling: Established Concepts and Emerging Topics. *International Journal of Molecular Sciences* 22, (2021).
9. Amarasekara, D. S., Kim, S. & Rho, J. Regulation of Osteoblast Differentiation by Cytokine Networks. *International Journal of Molecular Sciences* 22, (2021).

10. St Croix, C. M., Shand, S. H. & Watkins, S. C. Confocal microscopy: comparisons, applications, and problems. *Biotechniques* 39, (2005).
11. Turner, J., Lasek, S. & Szarowski, D. Confocal Optical Microscopy. in *Encyclopedia of Materials: Science and Technology* 1504–1509 (2001).
12. Mckinnon, K. M. Flow Cytometry: An Overview. *Current Protocols in Immunology* 120, 1–11 (2019).
13. Mahmood, T. & Yang, P. C. Western Blot: Technique, Theory, and Trouble Shooting. *North American Journal of Medical Sciences* 4, 429 (2012).
14. Lee, J.-M. *et al.* The effect of biomechanical stimulation on osteoblast differentiation of human jaw periosteum-derived stem cells. *Maxillofacial Plastic and Reconstructive Surgery* 39, (2017).
15. Bednarska, K., Klink, M. & Sulowska, Z. Application of intracellular alkaline phosphatase activity measurement in detection of neutrophil adherence in vitro. *Mediators of Inflammation* 2006, 1–8 (2006).
16. Adams, G. A beginner's guide to RT-PCR, qPCR and RT-qPCR. *The Biochemist* 42, 48–53 (2020).
17. Gregory, C. A., McNeill, E. P. & Pan, S. Preparation of osteogenic matrices from cultured cells. in *Methods in Cell Biology* vol. 156 15–43 (Academic Press Inc., 2020).
18. Hartung, A. *et al.* Different Routes of Bone Morphogenic Protein (BMP) Receptor Endocytosis Influence BMP Signaling. *Molecular and Cellular Biology* 26, 7791–7805 (2006).
19. Naslavsky, N. & Caplan, S. The enigmatic endosome-sorting the ins and outs of endocytic trafficking. *Journal of Cell Science* 131, (2018).

20. Grant, B. D. & Donaldson, J. G. Pathways and mechanisms of endocytic recycling. *Nat Rev Mol Cell Biol* 10, 608 (2009).
21. Ge, J. *et al.* The Size of Mesenchymal Stem Cells is a Significant Cause of Vascular Obstructions and Stroke. *Stem Cell Reviews and Reports* 10, 295–303 (2014).
22. Lennon, D. P., Schluchter, M. D. & Caplan, A. I. The Effect of Extended First Passage Culture on the Proliferation and Differentiation of Human Marrow-Derived Mesenchymal Stem Cells. *Stem Cells Translational Medicine* 1, 279–288 (2012).
23. Kwist C Bridges K J L Burg, K. W. The effect of cell passage number on osteogenic and adipogenic characteristics of D1 cells. *Cytotechnology* 68, 1661–1667 (2016).
24. Fink, S. L. & Cookson, B. T. Apoptosis, pyroptosis, and necrosis: Mechanistic description of dead and dying eukaryotic cells. *Infection and Immunity* vol. 73 1907–1916 Preprint at <https://doi.org/10.1128/IAI.73.4.1907-1916.2005> (2005).
25. Kokabua, S., Katagiria, T., Yodab, T. & Rosenc, V. Role of Smad phosphatases in BMP-Smad signaling axis-induced osteoblast differentiation. *Journal of Oral Biosciences* vol. 54 73–78 Preprint at <https://doi.org/10.1016/j.job.2012.02.003> (2012).
26. Sikand, K., Singh, J., Ebron, J. S. & Shukla, G. C. Housekeeping Gene Selection Advisory: Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) and β -Actin Are Targets of miR-644a. *PLoS ONE* 7, (2012).
27. Aubin, J. Regulation Of Osteoblast Formation. *Reviews in Endocrine and Metabolic Disorders* 2, 81–94 (2001).
28. Köllmer, M., Buhrman, J. S., Zhang, Y. & Gemeinhart, R. A. Markers Are Shared Between Adipogenic and Osteogenic Differentiated Mesenchymal Stem Cells. *Journal of Developmental Biology and Tissue Engineering* 5, 18–25 (2013).

29. Bruderer, M., Richards, R. G., Alini, M. & Stoddart, M. J. Role and regulation of runx2 in osteogenesis. *European Cells and Materials* 28, 269–286 (2014).
30. Anderson, H. C. Vesicles associated with calcification in the matrix of epiphyseal cartilage. *Journal of Cell Biology* 41, 59–72 (1969).
31. Chan, A. K. C., Heathman, T. R. J., Coopman, K. & Hewitt, C. J. Multiparameter flow cytometry for the characterisation of extracellular markers on human mesenchymal stem cells. *Biotechnology Letters* 36, 731–741 (2014).
32. Granéli, C. *et al.* Novel markers of osteogenic and adipogenic differentiation of human bone marrow stromal cells identified using a quantitative proteomics approach. *Stem Cell Research* 12, 153–165 (2014).
33. Ho, K. H. & Patrizi, A. Assessment of common housekeeping genes as reference for gene expression studies using RT-qPCR in mouse choroid plexus. *Scientific Reports* 11, (2021).

CHAPTER 6. CONCLUSIONS AND FUTURE PROJECT OUTLOOK

6.1 OVERVIEW

The general aim of this project was to provide the groundwork necessary for bridging the gap between synthetic biology and tissue engineering. Due to the current problems associated with growth factor therapeutics, including short half-life, poor dose retention and large dose administration, the overall concept of the project centered around on-demand production of growth factor capable of inducing osteogenesis differentiation of human mesenchymal stem cells (hMSCs) to improve bone regeneration properties in the case of bone fractures or osteoporosis. To advance towards this goal, growth factor-reporter protein fusion, compatible with a cell-free expression system (CFES) was firstly designed. The cell-free production of the chimera consisting of bone morphogenetic protein 2 (BMP2) and a reporter protein deGFP was measured across various conditions to determine optimal yield. The effectiveness of incorporating the working cell-free expression system into agar, agarose and sodium alginate/Pluronic F127 was briefly investigated. Next, the chimera was expressed in *E.coli* for characterisation purposes, including the effect of conjugating the growth factor with a reporter protein on their physical properties. Finally, the bioactivity of the generated protein was explored with mesenchymal stem cells to measure the osteogenic differentiation potential of the system.

In Chapter 2, several plasmids compatible with the cell-free expression system were designed to encode for a growth factor, vascular endothelial growth factor (VEGF) or BMP2, conjugated to a reporter protein including EGFP, mCherry and deGFP. The pCellFree plasmids which contain EGFP or mCherry at either N- or C-terminal of the multiple cloning site were successfully modified by addition of BMP2 or VEGF sequence using gateway cloning. Next, restriction digestion was used as a method to create pTar-

BMP2-EGFP for a two-plasmid system. The promoter of this plasmid is only recognised by a sigma factor 28 which had to be additionally supplemented. Finally, Gateway cloning was utilised to successfully build pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500.

Chapter 3 presented the data from cell-free protein expression using the series of constructed plasmids as input DNA. The production yields were compared using fluorescence output. This method of monitoring was made possible due to the fluorescent chromophore being synthesised and matured as part of the fusion protein. It was established that certain pCellFree plasmids were expressed in PURExpress, a reconstituted protein synthesis system, as well as the in-house *E.coli* CFES, however, only in the presence of Pr1-T7RNAP plasmid. This demonstrated that the transcription of pCellFree plasmids was driven by an optimal amount of a T7 RNA polymerase. A two-plasmid system of pTar-BMP2-EGFP and pBEST-p15a-OR2-OR1-Pr-UTR1-sigma28-T500 also generated a positive signal of fluorescence with in-house CFES, when correct ratios of the plasmids were conserved. It also provided another excellent example of the ability to implement further transcription regulation with the use of multiple plasmids. The highest cell-free protein producing plasmid that was generated in this thesis was the pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500, synthesising around 12.4 μM of the chimera per reaction. It was determined that the optimal input DNA concentration for this plasmid was at 3 nM before reaching a threshold. The reaction yields were determined for both the human physiological temperature of 37°C and the plasmid growth condition temperature of 29°C. Finally, CFES with pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500 was incorporated into agar, agarose and the bioink (sodium alginate and Pluronic F127). These demonstrated a positive fluorescence signal, indicating successful production of the chimera. Higher fluorescence signal was obtained from agar and agarose CFES gels than the bioink CFES, however, large error bars show the need for more consistent methods of generating the CFES gels. Nonetheless, for the first time to the best of the

author's knowledge, the cell-free production of a growth factor within gels was demonstrated, providing advancements for generation of smart materials.

The highly successful cell-free results with pBEST-OR2-OR1-Pr-UTR1-deGFP-BMP2-T500 led to further investigation of this particular chimera. In Chapter 4, this plasmid was transformed into *E.coli* cells for recombinant production and subsequent characterisation. Due to a high number of disulphide bonds in the chimera, presence of inclusion bodies was detected, requiring the process of protein refolding. Subsequently, after IMAC and SEC purifications, the resulting chimera consisted of a monomeric and dimeric mixture. Circular dichroism has shown the chimera to have a high β -sheet content in its secondary structure, as expected from its primarily β -sheet constituents, BMP2 and deGFP. The tertiary structure of the chimera was more challenging to elucidate due to the monomeric and dimeric presence in the sample. However, it was possible to determine that the conjugation between BMP2 and deGFP did not have a significant impact on the fluorescence properties. Dynamic light scattering revealed a large hydrodynamic radius difference between EGFP at 3.87 ± 0.03 nm and the chimera at 16.94 ± 0.13 nm, suggesting that the dimerisation brings together two BMP2 monomers, each attached with a deGFP molecule, as predicted.

In Chapter 5, the bioactivity of the chimera was explored by utilising numerous assays. Firstly, deGFP-BMP2 interactions with human mesenchymal stems cells (hMSCs) were observed using confocal microscopy, confirming the uptake of the chimera over time. Flow cytometry further validated the evidence of these protein-cell interactions, as hMSCs incubated with the *E.coli*-produced chimera displayed higher fluorescence than EGFP interactions. Downstream signalling initiated by BMP2 was apparent by conducting Western blotting, where both *E.coli*- and CFES deGFP-BMP2 generated SMAD 1/5/8 phosphorylation, a hallmark of BMP2 binding to cell-surface receptors. Subsequential osteogenic differentiation triggers were successfully delineated by measuring alkaline phosphatase (ALP) activity, osteogenic markers gene regulation and calcium deposition

by osteoblasts. ALP activity was found to be significantly heightened in hMSCs cultured with either the *E.coli*-produced chimera and CFES reaction with the produced chimera. It was shown that ALP activity can be increased by the chimera alone, with no other supplements in the expansion media, implying strong osteogenic potency of deGFP-BMP2. At a two-week timepoint of cell culture, early differentiation markers *Col1A1* and *ALP*, and a stable osteogenic marker *RUNX2* were upregulated in cells with osteogenic media containing *E.coli*-made chimera or CFES reaction containing the chimera. In contrast, a late osteogenic marker, *BGLAP*, was downregulated, commensurate with the progress of osteogenic differentiation. Finally, calcium deposits were confirmed at a four-week cell culture mark, by performing Alizarin Red S staining. This corresponded to the matrix mineralisation which is a process occurring during late bone remodelling. Together these results showed osteogenic differentiation commitment capability of the chimera, which surpassed that of a positive control, rhBMP2, in some assays.

To continue towards the main goal of on-demand BMP2 production from an injectable gel located at a fracture site, there are a few paths which have not been fully explored here and would be interesting to examine. This includes a method of protein synthesis regulation as it can be useful to mimic the successive and sustained release of signals during bone regeneration. One approach to consider is the use of genetic logic gates which would provide a tight control over the amount of growth factors produced each time. A system with activating and repressing genes could be employed for circuit design¹. Another method could utilise a Light-Oxygen-Voltage (LOV) sensing domain flavo-protein. This family of proteins are currently used as an optogenetics tool that generates a cellular response upon illumination with 450nm blue light². The use of optogenetic switches *in vitro* has now been well established³. However, if blue light stimulation is to be used *in vivo*, the penetration depth might prevent sufficient gene expression. Other examples that could be considered include ultrasound burstable capsules containing CFES. This could be built upon the idea by Kennedy *et al.*, where sequential release of nanoparticle payloads

Chapter 6 - Conclusions and Future Project Outlook

from alginate capsules would be released upon the application of ultrasound⁴. If this approach was applied to the CFES capsules, upon stimulation, the capsules would burst to release the growth factors and facilitate the differentiation of mesenchymal stem cells.

Ultimately, however, the array of plasmids, cell-free expression systems, and physical chassis for CFES tested in this thesis vastly increase the possibilities of alternative regulation and release mechanisms, lending credence to the future regenerative medicine applications of CFES-generated growth factor for stimulated osteogenesis differentiation of hMSCs.

6.1 BIBLIOGRAPHY

1. Wang, B., Kitney, R. I., Joly, N. & Buck, M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nature Communications* 2, 508 (2011).
2. Pudasaini, A., El-Arab, K. K. & Zoltowski, B. D. LOV-based optogenetic devices: light-driven modules to impart photoregulated control of cellular signaling. *Front Mol Biosci* 2, 18 (2015).
3. Polstein, L. R., Juhas, M., Hanna, G., Bursac, N. & Gersbach, C. A. An Engineered Optogenetic Switch for Spatiotemporal Control of Gene Expression, Cell Differentiation, and Tissue Morphogenesis. *ACS Synthetic Biology* 6, 2003–2013 (2017).
4. Kennedy, S. *et al.* Sequential release of nanoparticle payloads from ultrasonically burstable capsules. *Biomaterials* 75, 91–101 (2016).

A. APPENDIX A

GATEWAY CLONING OF pCELLFREE

Nucleotide sequence of VEGF with attB1 and attB2 sequences in red: (627bp)

ACA AGT TTG TAC AAA AAA GCA GGC TTC ATG AAC TTT CTG CTG TCT
TGG GTG CAT TGG AGC CTT GCC TTG CTG CTC TAC CTC CAC CAT GCC
AAG TGG TCC CAG GCT GCA CCC ATG GCA GAA GGA GGA GGG CAG AAT
CAT CAC GAA GTG GTG AAG TTC ATG GAT GTC TAT CAG CGC AGC TAC
TGC CAT CCA ATC GAG ACC CTG GTG GAC ATC TTC CAG GAG TAC CCT
GAT GAG ATC GAG TAC ATC TTC AAG CCA TCC TGT GTG CCC CTG ATG
CGA TGC GGG GGC TGC TGC AAT GAC GAG GGC CTG GAG TGT GTG CCC
ACT GAG GAG TCC AAC ATC ACC ATG CAG ATT ATG CGG ATC AAA CCT
CAC CAA GGC CAG CAC ATA GGA GAG ATG AGC TTC CTA CAG CAC AAC
AAA TGT GAA TGC AGA CCA AAG AAA GAT AGA GCA AGA CAA GAA AAT
CCC TGT GGG CCT TGC TCA GAG CGG AGA AAG CAT TTG TTT GTA CAA
GAT CCG CAG ACG TGT AAA TGT TCC TGC AAA AAC ACA GAC TCG CGT
TGC AAG GCG AGG CAG CTT GAG TTA AAC GAA CGT ACT TGC AGA TGT
GAC AAG CCG AGG CGG GAC CCA GCT TTC TTG TAC AAA GTG GTC

Nucleotide sequence of full length BMP2 with attB1 and attB2 sequences in red: (1242bp)

ACA AGT TTG TAC AAA AAA GCA GGC TTC ATG GTG GCC GGG ACC CGC
TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC CTC CTG GGC GGC GCG
GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG TTC GCG GCG GCG
TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG GTC CTG AGC
GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA CAG AGA
CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA GAC
CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC
CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC
TTC CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG
AAA ACA ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG
GAG GAG TTT ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG
ATG CAA GAT GCT TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT
AAT ATT TAT GAA ATC ATA AAA CCT GCA ACA GCC AAC TCG AAA TTC

CCC GTG ACC AGA CTT TTG GAC ACC AGG TTG GTG AAT CAG AAT GCA
AGC AGG TGG GAA AGT TTT GAT GTC ACC CCC GCT GTG ATG CGG TGG
ACT GCA CAG GGA CAC GCC AAC CAT GGA TTC GTG GTG GAA GTG GCC
CAC TTG GAG GAG AAA CAA GGT GTC TCC AAG AGA CAT GTT AGG ATA
AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC TGG TCA CAG ATA AGG
CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA GGG CAT CCT CTC
CAC AAA AGA GAA AAA CGT CAA GCC AAA CAC AAA CAG CGG AAA CGC
CTT AAG TCC AGC TGT AAG AGA CAC CCT TTG TAC GTG GAC TTC AGT
GAC GTG GGG TGG AAT GAC TGG ATT GTG GCT CCC CCG GGG TAT CAC
GCC TTT TAC TGC CAC GGA GAA TGC CCT TTT CCT CTG GCT GAT CAT
CTG AAC TCC ACT AAT CAT GCC ATT GTT CAG ACG TTG GTC AAC TCT
GTT AAC TCT AAG ATT CCT AAG GCA TGC TGT GTC CCG ACA GAA CTC
AGT GCT ATC TCG ATG CTG TAC CTT GAC GAG AAT GAA AAG GTT GTA
TTA AAG AAC TAT CAG GAC ATG GTT GTG GAG GGT TGT GGG TGT CGC
GAC CCA GCT TTC TTG TAC AAA GTG GTC

Primers based on attB sites for linearisation of VEGF and BMP2 plasmids:

Forward primer sequence (5'-3'): ACA AGT TTG TAC AAA AAA GCA GGC TTC

Reverse primer sequence (5'-3'): GAC CAC TTT GTA CAA GAA AGC TGG GTC

pCellFree plasmids:

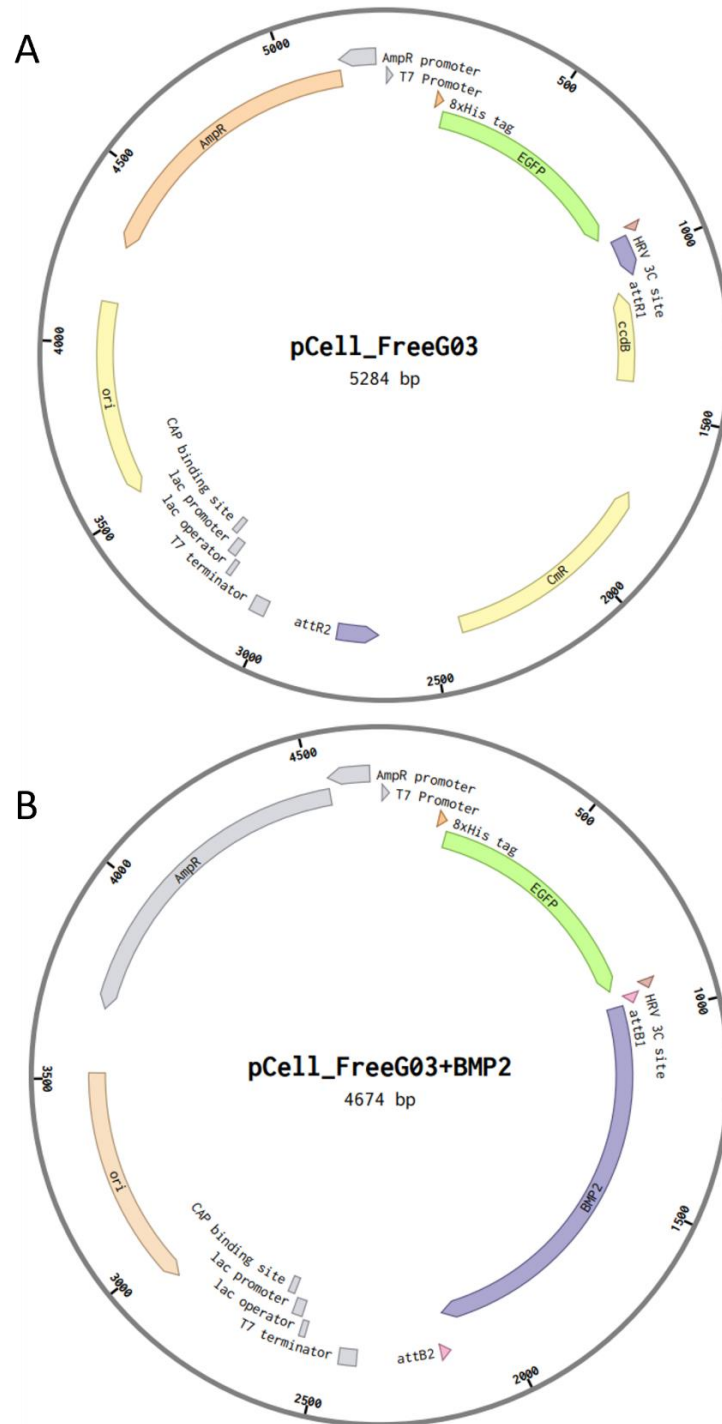


Figure A.1 Maps of pCellFreeG03 plasmids. A: Unmodified pCellFreeG03 consisting of EGFP with 8xHis tag at the N-terminal and HRV-3C site along with attB sites at the C-terminal. B: Gateway cloning modified pCellFreeG03 consisting of 8xHis tag, EGFP, HRV-3C site, full length BMP2, respectively.

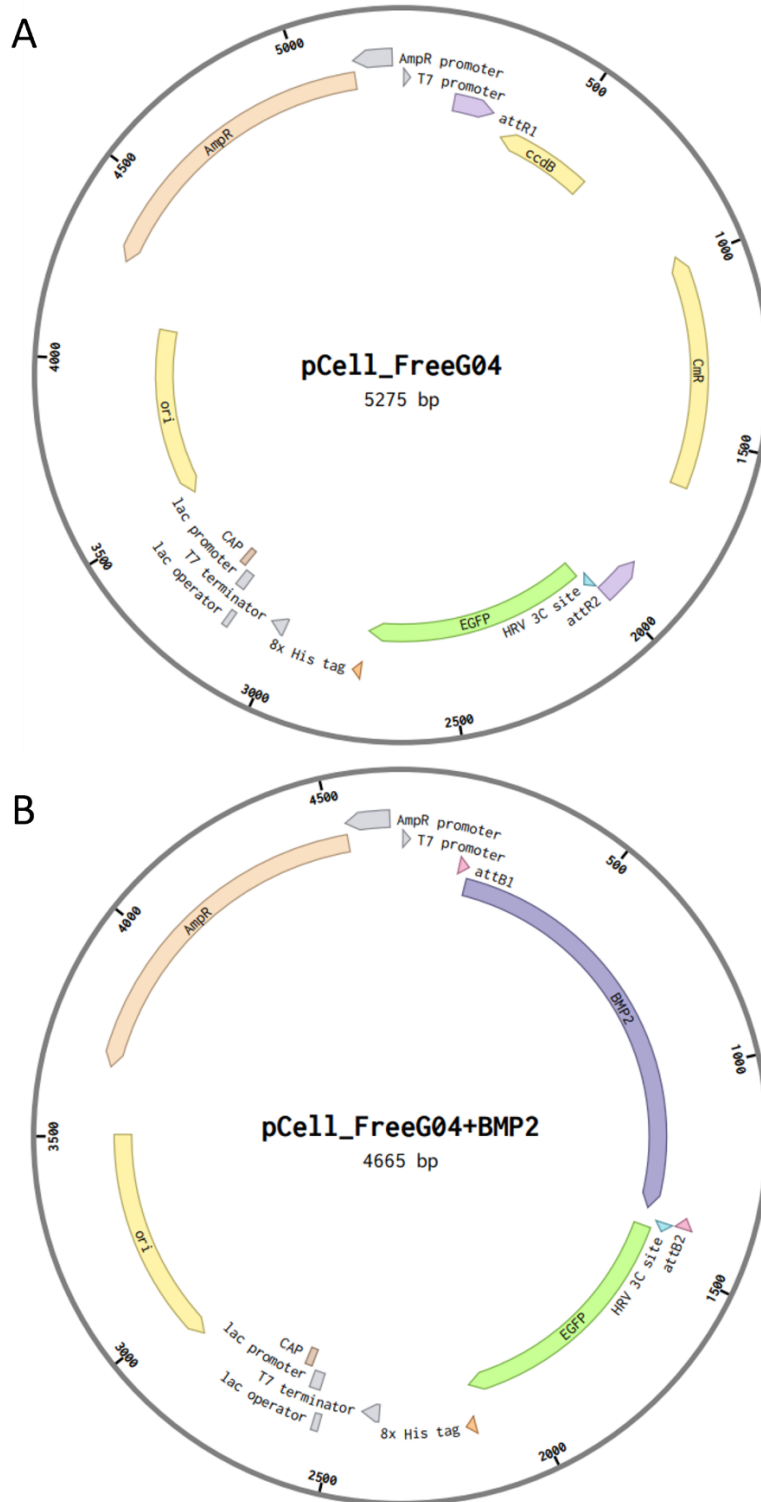


Figure A.2 Maps of pCellFreeG04 plasmids. A: Unmodified pCellFreeG04 consisting of EGFP with 8xHis tag at the C-terminal and HRV-3C site along with attB sites at the N-terminal. B: Gateway cloning modified pCellFreeG04 consisting of full length BMP2, HRV-3C site, EGFP and 8xHis tag, respectively.

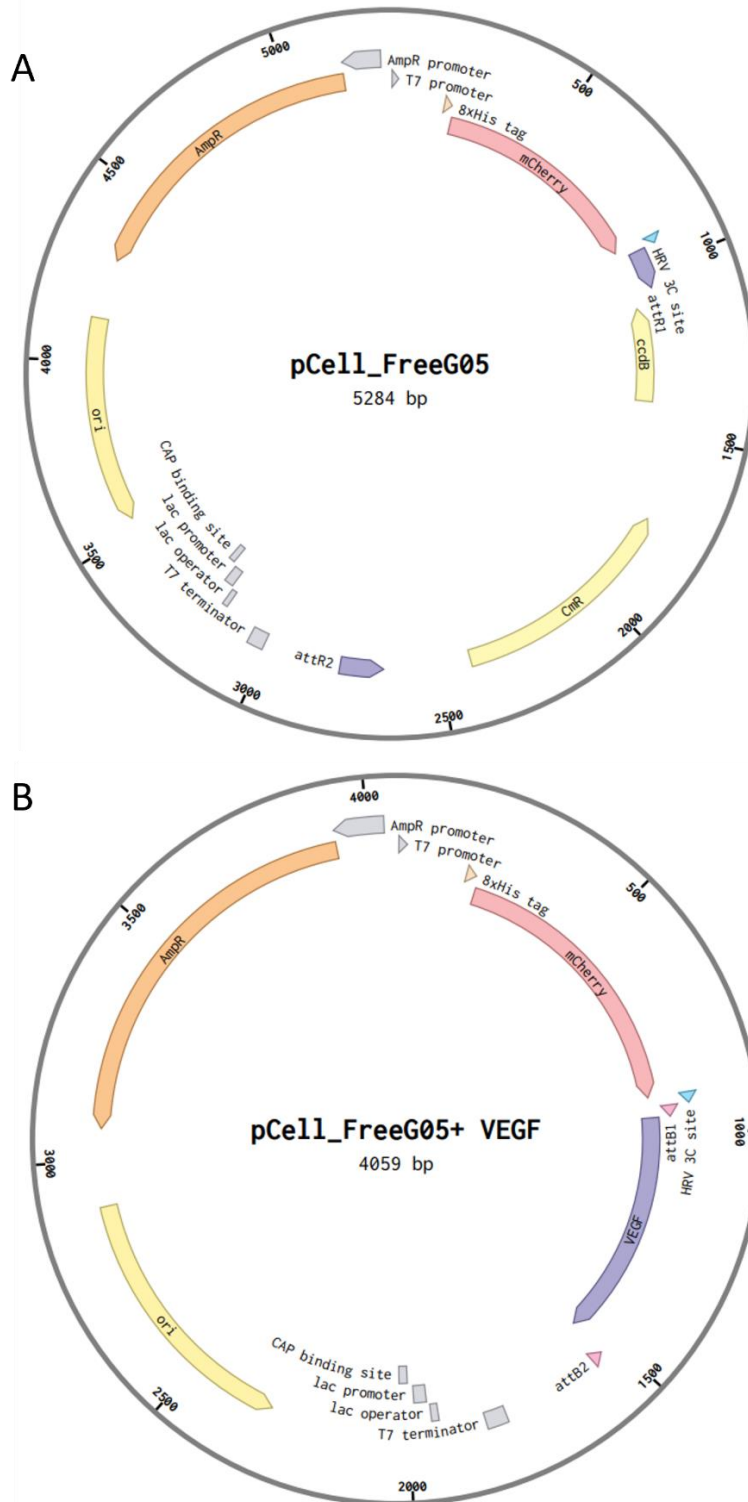


Figure A.3 Maps of pCellFreeG05 plasmids. A: Unmodified pCellFreeG05 consisting of mCherry with 8xHis tag at the N-terminal and HRV-3C site along with attB sites at the C-terminal. B: Gateway cloning modified pCellFreeG05 consisting of 8xHis tag, mCherry, HRV-3C site, VEGF, respectively.

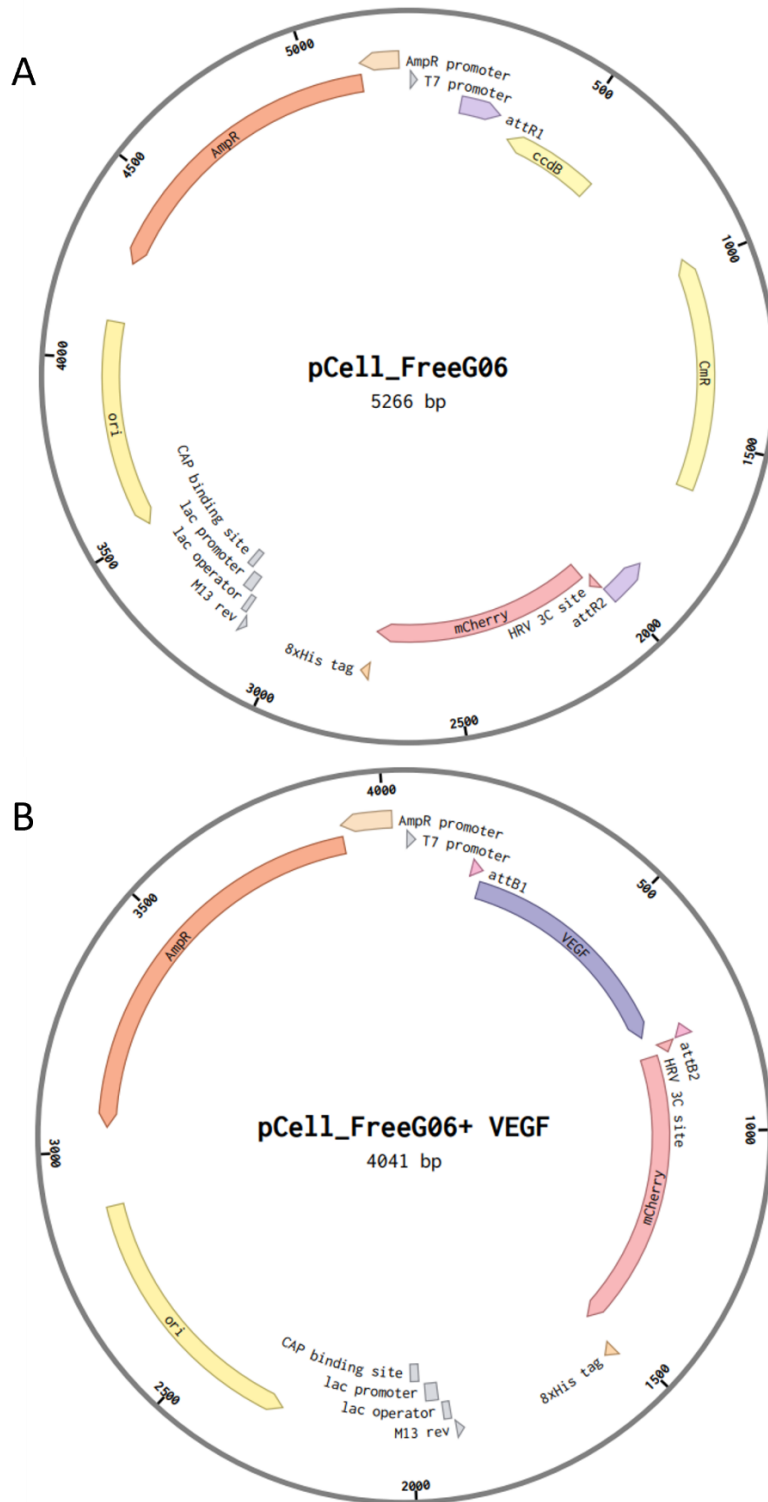


Figure A.4 Maps of pCellFreeG06 plasmids. A: Unmodified pCellFreeG06 consisting of mCherry with 8xHis tag at the C-terminal and HRV-3C site along with attB sites at the N-terminal. B: Gateway cloning modified pCellFreeG06 consisting of VEGF, HRV-3C site, mCherry and 8xHis tag, respectively.

Pr1T7 plasmid:

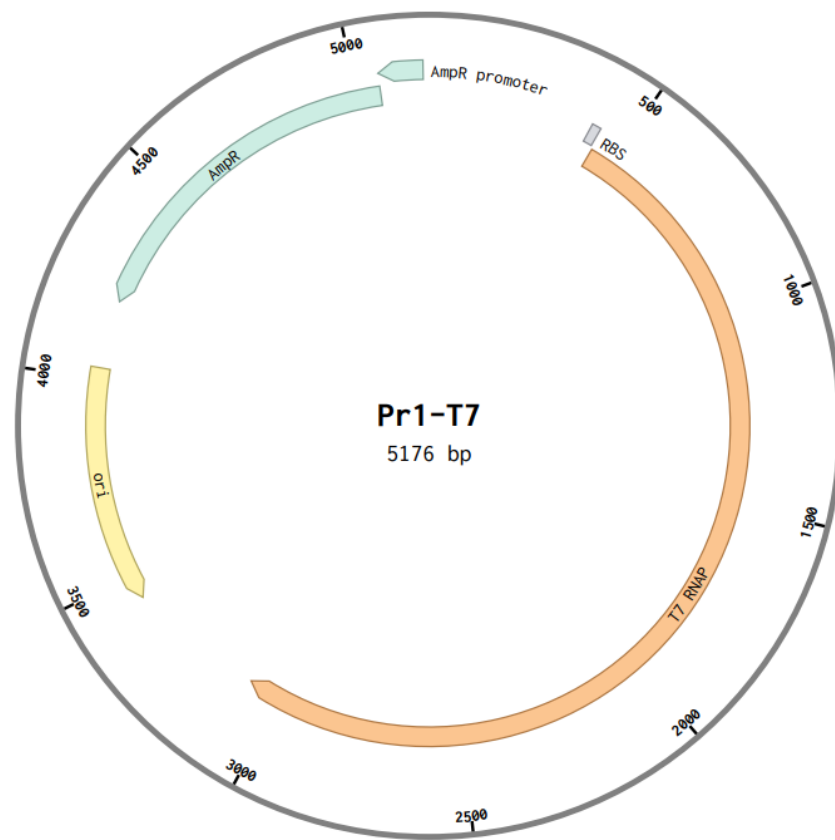


Figure A.5 Map of Pr1-T7 plasmid. Encodes for a T7 RNA polymerase. It was used in conjunction with pCellFree plasmids to bring about their transcription.

ENGINEERING A TWO-PLASMID SYSTEM

Mature-length BMP2, E.coli optimized nucleotides in pEX-A128 vector backbone. TEV cleavage site in blue: (380bp)

```
ATG CAG GCC AAA CAT AAA CAG CGT AAA CGT CTG AAA AGC AGC
TGC AAA CGT CAT CCG CTG TAT GTT GAT TTT AGT GAT GTT GGT
TGG AAC GAT TGG ATT GTT GCA CCG CCT GGT TAT CAT GCA TTT
TAT TGT CAT GGT GAA TGT CCG TTT CCG CTG GCA GAT CAT CTG
AAT AGC ACC AAT CAT GCA ATT GTT CAG ACC CTG GTT AAT AGC
GTG AAT AGC AAA ATT CCG AAA GCA TGT TGT GTT CCG ACC GAA
CTG AGC GCA ATT AGC ATG CTG TAT CTG GAT GAA AAT GAA AAG
GTG GTG CTG AAA AAC TAT CAG GAT ATG GTT GTT GAA GGT TGT
GGT TGT CGT GAA AAT CTG TAT TTT CAG AGC
```

Sigma28 plasmid:

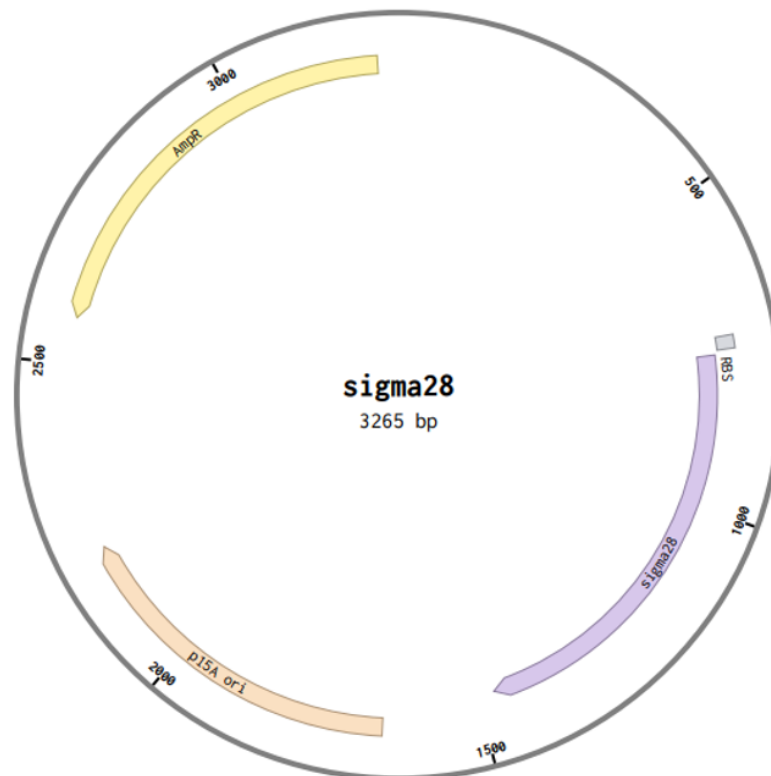


Figure A.6 Map of sigma 28 plasmid. Once expressed, it enables transcription of plasmids with pTar promoters.

AqpZ expression plasmid:

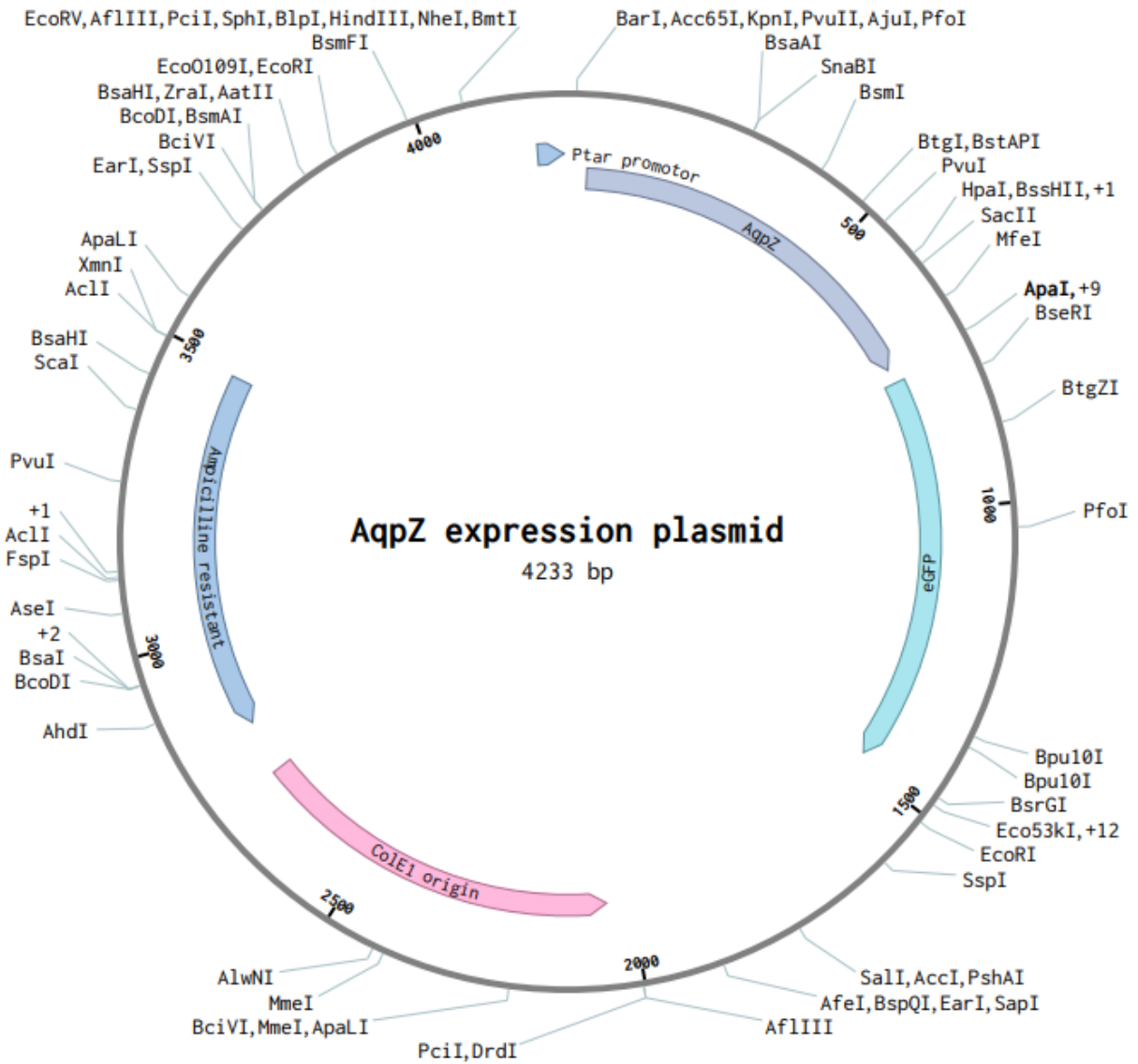


Figure A.7 Map of unmodified expression plasmid. Requires sigma factor 28 to be transcribed. Encodes for AqpZ with EGFP at the C-terminal. This plasmid was used to replace AqpZ gene with mature-length BMP2 gene.

BMP2-eGFP expression plasmid:

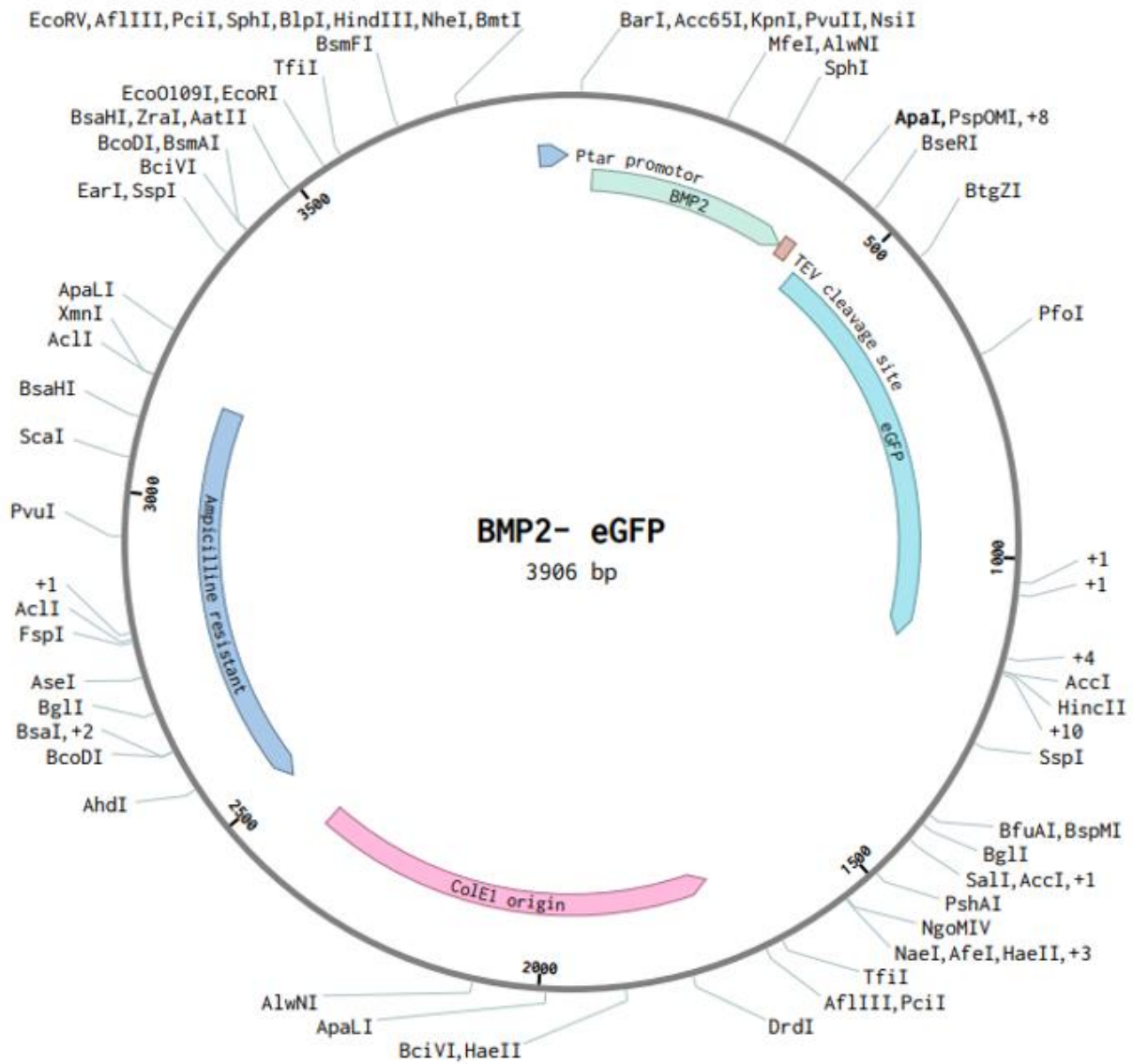


Figure A.8 Map of modified expression plasmid encoding for mature-length BMP2, TEV cleavage site and EGFP, respectively. Requires sigma factor 28 to be transcribed.

GIBSON CLONING OF DEGFP-BMP2

Primers used to linearise and generate overhangs for the BMP2 fragment:

Forward sequence (5'-3'): GCC GGG ATC GAA AAT CTG TAT TTT CAG AGC CAG
GCC AAA CAT AAA CAG CGT

Reverse sequence (5'-3'): GCT TTG CTC GAG TTA GTG GTG GTG GTG GTG GTG
ACG ACA ACC ACA ACC TTC AAC

Primers used to linearise and generate overhangs for the vector fragment:

Forward sequence (5'-3'): GGT TGT CGT CAC CAC CAC CAC CAC CAC TAA CTC
GAG CAA AGC CCG CCG AAA G

Reverse sequence (5'-3'): ACG ACA ACC GCT CTG AAA ATA CAG ATT TTC GAT
CCC GGC GGC GGT CAC

BMP2 sequence used here was from pEX-A128 vector. The same as in the two-plasmid system.

deGFP plasmid:

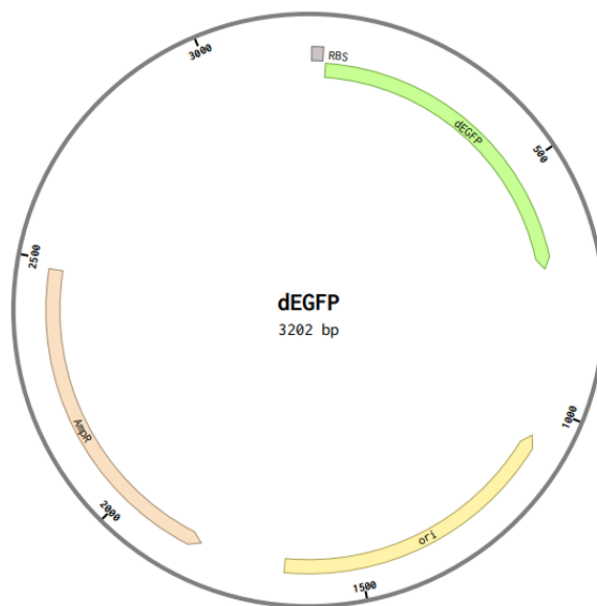


Figure A.9 Map of pBEST-OR2-OR1-Pr-UTR1-deGFP-T500.

deGFP-BMP2 plasmid:

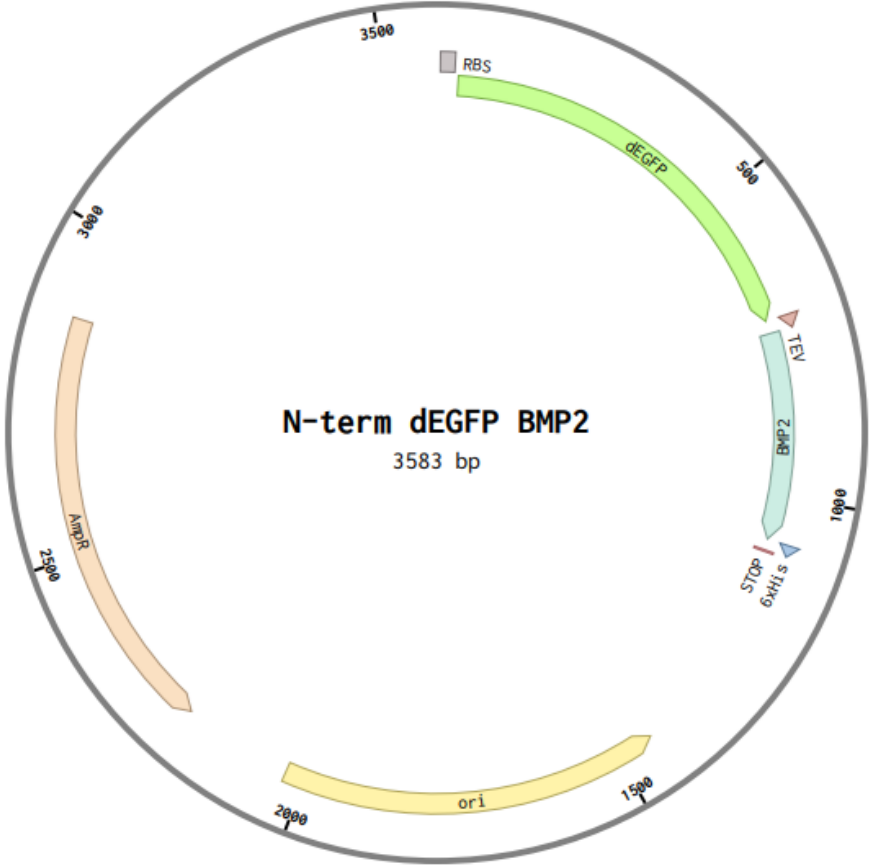


Figure A.10 Map of modified pBEST-OR2-OR1-Pr-UTR1-deGFP-T500, now encoding for deGFP, TEV cleavage site, mature-length BMP2, 6xHis tag, respectively.

B. APPENDIX B

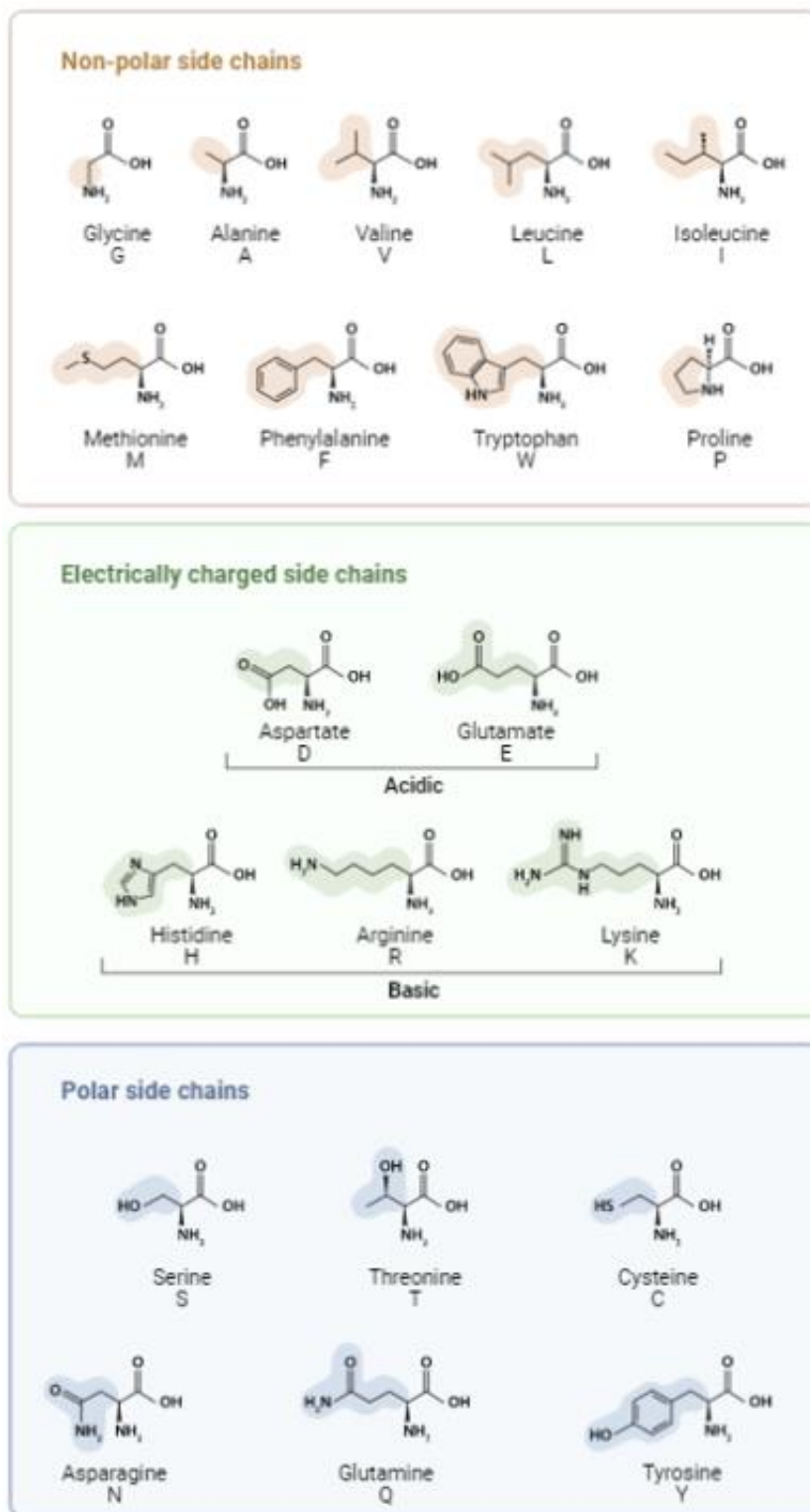


Figure B.1 Naturally occurring amino acids. Generated from biorender.com.

C. APPENDIX C

Amino acid sequence of deGFP-BMP2:

MELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLPVPWP
TLVTTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAE
VKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKI
RHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDH MVLLE
FVTAAGIENLYFQSQA KHKQRKRLKSSCKRHPLYVDFSDVGWNDWIVAPPGYHA
FYCHGECFPPLADHLNSTNHAIVQTLVNSVNSKIPKACCVPTLSAISMLYLDENE
KVV LKNYQDMVVEGCGCRHHHHHHH

Green= deGFP

Brown= TEV cleavage site

Blue= mature-length BMP2

Black= 6xHis tag

Amino acid sequence of EGFP:

MVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTCLKFICTTGKLP
VPWPTLVTTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIK
VNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALS KDPNEKRDH
MVLLEFVTAAGITLGMDELYK