Investigation of Bacterial Cellulose Production in Genetically Modified *Escherichia coli*

Gizem Buldum

Department of Chemical Engineering

Imperial College London

A thesis submitted for the degree of Doctor of Philosophy and the Diploma of Imperial College London,

2016 February
Declaration of Originality

The work described in this thesis was carried out by author in the Department of Chemical Engineering, Imperial College London, United Kingdom between October 2011 and October 2015 under the supervision of Prof. Athanasios Mantalaris and Prof. Alexander Bismarck. Except where acknowledge, the material is the original work of the author and includes nothing, which is the outcome of work in collaboration, and no part of it, has been submitted for a degree at any other university.

Gizem Buldum
Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.
Acknowledgements

This PhD research has been quite important journey for my life, punctuated by extraordinary experiences and exciting knowledge with the contribution of great number of people. It is a pleasure to thank every one of them for supporting me on this journey.

First of all, I would like to express my deep and sincere gratitude to Prof. Athanasios Mantalaris, my academic supervisor, for his valuable guidance, everlasting encouragement and support. Also, I would like to acknowledge Prof. Mantalaris for the great freedom he has given me throughout my research, but also for the hard question which incented me to widen my research from various perspectives. I would also like to express my deep appreciation to my co-supervisor Prof. Alexander Bismarck for sharing his diverse experiences, valuable guidance and advices throughout my research. It was a great pleasure working both of them.

I would like to thank The Graduate Studies Abroad Programme by Ministry of National Education of Turkey for financial support.

I’m also grateful to my seniors, Dr. Koon-yang Lee and Dr. Lia Fauzi for their help and guidance through my research.

I would also thank Sarah Payne and Susi Underwood for their patience and help for my numerous requests over the years.

I’ve been very fortune sharing the same office with Joana Dos Santos, Susana Dos Santos, Mauricio Zamorano and Osama Fallata. I would like to thank all of them for the nice and fun atmosphere and making easier this period for me. I would like to thank my colleague and also flatmate Ana Quiroga for supporting me all my hard and good times. I’m thankful to my fellow friends Keeran, Chonlatep and Mark for their great friendship and positive energy. I am also grateful to Iro Tsipa for being a great labmate with constructive exchanges. My special thanks are to Maria
Papathanasiou and Mariano Latorre, for sharing all my good and hard times during this journey with their everlasting support and great understanding.

Finally, my deepest gratitude is to my dearest parents, Belgin Buldum and Zeynel Abidin Buldum, for encouraging and supporting me. Since I left my home country for this long academic journey, I felt them always next to me with their unconditional support, selfless love and sacrifice. This thesis is dedicated to my parents.
ABSTRACT

Cellulose is the major biopolymer on earth with tremendous economic importance as it has been utilised in a multitude of industrial applications including tissue-engineering products, composite materials and electronics. It exhibits outstanding physical and mechanical properties when compared to plant-based cellulose.

Although *A. xylinum* is the most efficient producer of bacterial cellulose (BC), its long doubling time results in insufficient yields and high cost. In this study, a novel and functional BC production system was developed by recombinant DNA technology. The simultaneous expression of bacterial cellulose synthase operon (*bcsABCD*) and its upstream operon (*cmcax* and *ccpAx*) was achieved by pBCS and pCDF, respectively. Three different *Escherichia coli* strains were utilised as host microorganisms: *E. coli* BL21 (DE3), *E. coli* HMS174 (DE3) and *E. coli* C41 (DE3). It was verified that *bcsABCD* and the upstream operon were successfully cloned and expressed in *E. coli* strains. Fermentation of genetically modified (GM) strains was conducted at various IPTG concentrations (0.025, 0.05, 0.1, 0.2, 0.5 and 1.0 mM) and various temperatures (22, 30, and 37 °C). BC production was achieved by genetically modified *E. coli* HMS174 (DE3) in the presence of 0.025 mM IPTG at 22°C. GM *E. coli* C41 (DE3) accomplished the production when IPTG supplement was lower than 0.2 mM at 22°C or 30°C. The products were characterised by SEM and FTIR, which exhibited that morphology of product was stain-specific. Finally a dynamic mathematical model was developed to design a fed-batch system capturing characteristics incorporating acetate inhibition and cell death, which allowed predicting glucose consumption, acetate production and induction time for batch cultures, resulted in a volumetric productivity of 1.7 mg/L.h.

In conclusion, this thesis reports the development of a novel BC production system by creating valuable cellulose-producing *E. coli* strains, resulting in a reproducible and stable recombinant expression system for potential improvement of BC.
CONTENTS

1. GENERAL INTRODUCTION .............................................................................................. 19
   1.1. Overview ................................................................................................................ 20
   1.2. Motivation and Novelty ......................................................................................... 21
   1.3. Roadmap of the thesis ........................................................................................... 22

2. GENERAL BACKGROUND ................................................................................................ 24
   2.1. Cellulose Types ....................................................................................................... 25
       2.1.1. Plant Cellulose ................................................................................................ 25
       2.1.2. Bacterial Cellulose ....................................................................................... 26
   2.2. Advantages of Bacterial Cellulose over Plant Cellulose ......................................... 27
   2.3. Metabolic Pathway of Cellulose-Producing Bacteria ............................................. 29
   2.4. Genetic Pathway of Acetobacter ........................................................................... 32
   2.5. Fermentative Production of BC ............................................................................. 34
       2.5.1. Composition of Culture Media ........................................................................ 34
       2.5.2. Bioprocess Parameters .................................................................................. 37
       2.5.3. Influence of pH on BC Production ................................................................ 37
       2.5.4. Influence of Temperature on BC production ................................................ 38
       2.5.5. Influence of Oxygen on BC production ........................................................ 38
   2.6. Genetic Modification of Bacteria to Enhance BC Production .................................. 39
   2.7. Recombinant DNA Technology and Molecular Cloning ...................................... 41
   2.8. Recombinant protein expression in Escherichia coli by plasmid vectors .............. 42
   2.9. Conclusion .............................................................................................................. 43

3. AIMS AND OBJECTIVES ................................................................................................... 45
   3.1. Importance of Bacterial Cellulose for Industrial Applications ............................... 46
   3.2. Research Challenges ............................................................................................. 48
   3.3. Proposed Solution and Objectives of This Study ................................................... 50

4. MATERIALS AND METHODS ........................................................................................... 52
   4.1. MATERIALS ............................................................................................................. 53
       4.1.1. Bacterial Strains and Plasmids ....................................................................... 53
       4.1.2. Chemicals ........................................................................................................ 54
       4.1.3. Growth Media for Escherichia coli ................................................................ 54
       4.1.4. Growth Media for Acetobacter xylinum ATCC 53582 ...................................... 55
       4.1.5. Gels, Buffers, Solutions and Other Materials .................................................. 56
       4.1.6. Kits .................................................................................................................. 57
       4.1.7. Other Stocks and Solutions ............................................................................ 57
4.2. METHODS ............................................................................................................... 59
4.2.1. Sterilization ........................................................................................................... 59
4.2.2. Preparation of Bacterial Stocks ............................................................................. 59
4.2.3. Preparation of Preculture ...................................................................................... 59
4.2.4. Growth Conditions ............................................................................................... 60
4.2.5. Monitoring Cell Growth ......................................................................................... 60
4.2.6. Viable Cell Count and Plasmid Stability Test ......................................................... 61
4.2.7. Plasmid Transformation into E. coli Cells ............................................................ 62
4.2.8. Determination of Dry Cell Weight ......................................................................... 63
4.2.9. Extraction of Proteins ............................................................................................ 64
4.2.10. DNA Isolation ..................................................................................................... 68
4.2.11. RNA Isolation ..................................................................................................... 69
4.2.12. Polymerase Chain Reaction ............................................................................... 71
4.2.13. Primer Design ..................................................................................................... 74
4.2.14. Agarose Gel Electrophoresis ............................................................................. 75
4.2.15. Reverse Transcription Procedure ...................................................................... 76
4.2.16. Real Time q-PCR ............................................................................................... 78
4.2.17. Restriction Digest Test ....................................................................................... 80
4.2.18. FTIR Spectroscopy ............................................................................................. 81
4.2.19. Scanning Electron Microscopy .......................................................................... 82
4.2.20. Congo Binding Assay ......................................................................................... 82
4.2.21. Quantification of Bacterial Cellulose Production ............................................... 83
4.2.22. Determination of Free Carbon Source .............................................................. 83
4.2.23. Operation of Stirred-tank Bioreactor ................................................................ 88
4.2.24. Statistical Analysis ............................................................................................. 88

5. DESIGN OF THE RECOMBINANT EXPRESSION SYSTEM .................................................. 89
5.1. Introduction and Aim ............................................................................................... 90
5.2. Methodology ........................................................................................................... 92
5.3. The Selection of The Host Cell ............................................................................... 92
5.4. The Choice of The Plasmid ..................................................................................... 94
5.4.1. Major Features of Expression Vectors ............................................................... 95
5.4.2. Protein Co-expression and Duet Vectors ........................................................... 98
5.5. Design of the Expression System ........................................................................... 101
5.5.1. One Plasmid Approach ....................................................................................... 101
5.5.2. Two Plasmid Approach ...................................................................................... 103
5.6. Cloning strategy ................................................................................................... 109
5.7. Restriction Digest Test Results .............................................................................. 116
5.8. Plasmid Transformation into E. coli Cells ............................................................ 117
5.9. Discussion ............................................................................................................... 117
5.10. Conclusion ............................................................................................................ 120

6. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED E. COLI STRAIN 1: BL21 (DE3) ............................................................................................................. 121

6.1. Introduction ............................................................................................................ 122

6.1.1. E. coli BL21 (DE3) as a Host Microorganism .................................................. 122
6.1.2. Effects of Plasmid Presence on E. coli ............................................................ 123
6.1.3. T7 RNA Polymerase System ........................................................................ 125
6.1.4. Bacterial Growth Phases .............................................................................. 126
6.1.5. Aims and Objectives ...................................................................................... 127

6.2. Methodology ......................................................................................................... 128

6.3. Results .................................................................................................................. 130

6.3.1. Expression Systems ....................................................................................... 130
6.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of E. coli BL21 (DE3) .......................................................................................................................... 131
6.3.3. Gene Expression Analysis of BC Synthase Related Genes ........................ 147
6.3.4. Congo Binding Analysis ............................................................................... 151
6.3.5. SDS PAGE Analysis ..................................................................................... 152

6.4. Discussion ............................................................................................................. 154

6.5. Conclusion ............................................................................................................ 158

7. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED E. COLI STRAIN 2: HMS174 (DE3) ............................................................................................................. 159

7.1. Introduction ............................................................................................................ 160

7.1.1. E. coli HMS174 (DE3) as a Host Microorganism ........................................ 160
7.1.2. Effects of Plasmid Presence on E. coli HMS174 (DE3) ................................ 160
7.1.3. Cellulose production in E. coli ...................................................................... 161
7.1.4. Aims and Objectives ...................................................................................... 164

7.2. Methodology ......................................................................................................... 164

7.3. Results .................................................................................................................. 166

7.3.1. Expression Systems ....................................................................................... 166
7.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of E. coli HMS174 (DE3) .......................................................................................................................... 167
7.3.3. Gene Expression Analysis of BC Synthase Related Genes ........................ 177
7.3.4. Congo Binding Analysis ............................................................................... 178
7.3.5. SDS PAGE Analysis ..................................................................................... 180
7.3.6. Direct Detection of BC in Cultures ............................................................... 182
7.3.7. SEM Observation of BC ............................................................................... 183
7.3.8. ATR-FTIR Characterisation of Biosynthesised BC ..................................... 185

7.4. Discussion ............................................................................................................. 188
8. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED E. COLI STRAIN 3: C41 (DE3) ................................................................. 193

8.1. Introduction ........................................................................................................ 194

8.1.1. E. coli C41 (DE3) as a Host Microorganism ........................................ 194

8.1.2. Aims and Objectives ................................................................................ 196

8.2. Methodology .................................................................................................... 196

8.3. Results .............................................................................................................. 198

8.3.1. Expression Systems ................................................................................... 198

8.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of E. coli C41 (DE3) ... 199

8.3.3. Single plasmid approach ........................................................................... 206

8.3.4. Gene Expression Analysis of BC Synthase Related Genes ...................... 207

8.3.5. Direct Detection of BC in Cultures .............................................................. 209

8.3.6. SEM Observation of BC ............................................................................ 210

8.3.7. ATR-FTIR Characterisation of Biosynthesised BC ................................. 212

8.4. Discussion ....................................................................................................... 213

8.5. Conclusion ...................................................................................................... 216

9. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED E. COLI C41 (DE3) IN FED-BATCH SYSTEMS ................................................................. 217

9.1. Introduction .................................................................................................... 218

9.1.1. Limiting Factors in Batch Cultivation ....................................................... 218

9.1.2. Fed- batch Cultivation and Mathematical Models ................................... 222

9.1.3. Aims and Objectives ................................................................................. 223

9.2. Methodology .................................................................................................. 224

9.3. Results ........................................................................................................... 225

9.3.1. Comparison of Growth, Carbon Source Consumption, Acetate Production of E. coli in Batch Cultures .................................................. 225

9.3.2. Model structure ......................................................................................... 232

9.3.3. Model simulations .................................................................................... 239

9.3.4. Fed-Batch Cultures .................................................................................. 242

9.3.5. Microscope Images and Direct Observations of Product ....................... 247

9.4. Discussion ...................................................................................................... 251

9.5. Conclusion ..................................................................................................... 253

10. GENERAL CONCLUSIONS AND FUTURE WORK ................................................. 254

10.1. General Conclusions .................................................................................... 255

10.2. Future work .................................................................................................. 259
LIST OF FIGURES

Figure 1.1: Summary of the annual number of publications related to bacterial cellulose from the year 2000 to 2015 indexed within the Web of Science®. The keywords for the search included following terms: bacterial cellulose, microbial cellulose. ......................................................... 21
Figure 2.1: A schematic diagram showed structure (A) and composition (B) of a plant cell wall [8]. ........................................................................................................................................ 26
Figure 2.2: Glycosidic bond between β-D-glucose molecules (A) and hydrogen bonds on the same or the neighbour chain (B). ........................................................................................................ 27
Figure 2.3: Purification process of plant cellulose (A) and bacterial cellulose (B). .............. 30
Figure 4.1: Dilution of the cells on agar plate ........................................................................ 62
Figure 4.2: Schematic representation of chemical transformation ...................................... 63
Figure 4.3: Polymerase chain reaction .................................................................................. 72
Figure 4.4: Agarose gel electrophoresis. .............................................................................. 75
Figure 4.5: Reverse Transcription ...................................................................................... 77
Figure 4.6: Amplification plot ............................................................................................. 79
Figure 4.7: Instrumented-bioreactor equipped with feedback control-loop[110]. ............... 88
Figure 5.1: Major features of an expression vector. ............................................................. 94
Figure 5.2: Design of multiple cloning side based on one plasmid: pET-derivative. ............. 101
Figure 5.3: Amplification of target fragment (bcsABCD, cmcox, ccpAx). Lane 1: molecular weight marker, Lane 2: PCR product ........................................................................................................ 102
Figure 5.4: Open reading frames in a double-stranded DNA ............................................. 103
Figure 5.5: Original vector map and cloning regions of pETDuet-1.................................... 105
Figure 5.6: Original vector map and cloning regions of pCDFDuet-1. ................................. 106
Figure 5.7: Design of multiple cloning sides based on two plasmids: pBCS (A) and pCMP (B). ............................................................................................................................................. 108
Figure 5.8: Restriction sites (A) and zero-cutter enzymes (B) for cmcox. ......................... 111
Figure 5.9: Cloning strategy for pCMP ................................................................................ 114
Figure 5.10: Cloning strategy for pBCS ............................................................................. 115
Figure 5.11: Restriction digest samples for pCMP (A) and pBCS (B). For both gel images, Lane1: undigested, Lane 2: digested, Lane3 : DNA ladder. ........................................................................... 117
Figure 6.1: The T7 RNA Polymerase System .................................................................... 126
Figure 6.2: Bacterial Growth Phases .................................................................................. 127
Figure 6.3: Growth profiles of BL21 (DE3) cells at 37˚C ................................................... 133
Figure 6.4: Proportion of the plasmid-bearing BL21 (DE3) cells at 37˚C ............................ 134
Figure 6.5: Growth profiles of BL21 (DE3) cells at 37˚C ................................................... 135
Figure 6.6: Proportion of the plasmid-bearing BL21 (DE3) cells at 37˚C ......................... 136
Figure 6.7: Growth profiles of BL21 (DE3) cells at 30˚C ................................................... 138
Figure 6.8: Proportion of the plasmid-bearing BL21 (DE3) cells at 30˚C. ........................ 139
Figure 6.9: Growth profiles of BL21 (DE3) cells at 22˚C .................................................... 140
Figure 6.10: Proportion of the plasmid-bearing BL21 (DE3) cells at 22˚C ......................... 141
Figure 6.11: Growth profiles of BL21 (DE3) cells cultured in bioreactor/flask at 22˚C ....... 142
Figure 6.12: Proportion of the plasmid-bearing BL21 (DE3) cells cultured in bioreactor/flask at 22˚C. ............................................................................................................................................ 143
Figure 6.13: Growth profiles of BL21 (DE3) cells induced at OD=1.0, 1.5 and 2.0. .......... 145
Figure 6.14: Proportion of the plasmid-bearing BL21 (DE3) cells induced at OD=1.0, 1.5 and 2.0. ............................................................... 146
Figure 6.15: Relative gene expressions of BL21 (DE3) (pBCS + pCMP) after 3, 6, 18 hr induction by 0.2 mM IPTG 30˚C. ........................................ 148
Figure 6.16: Relative gene expressions of BL21 (DE3) (pBCS + pCMP), 15, 30, 60 min after induction by 0.2 mM IPTG at 30˚C. .................. 149
Figure 6.17: Relative gene expressions of BL21 (DE3) (pBCS + pCMP), 15, 30, 60 min after induction by 0.025 mM IPTG at 30˚C. .......... 150
Figure 6.18: Relative gene expressions of BL21 (DE3) (pBCS + pCMP), 18 hr after induction by 0.025 mM IPTG at 22˚C .................................. 150
Figure 6.19: Congo binding analysis of BL21 (DE3) (1) and BL21 (DE3) (pBCS + pCMP) (2) on agar plates.................................................. 151
Figure 6.20: Congo binding analysis of E. coli BL21 (DE3) (1) and E. coli BL21 (DE3) (pBCS + pCMP) (2) in liquid culture. ....................... 152
Figure 6.21: SDS-PAGE analysis of periplasmic (A) and insoluble (B) proteins extracted from E. coli BL21 (DE3) (pBCS + pCMP). ....................... 154

Figure 7.1: bcs operons and regulatory genes of cellulose biosynthesis in E. coli and A. xylinum. ................................................................. 163
Figure 7.2: Growth profiles of HMS174 (DE3) cells at 37˚C. ........................................ 169
Figure 7.3: Proportion of the plasmid-bearing HMS174 (DE3) cells at 37˚C. ............. 170
Figure 7.4: Growth profiles of HMS174 (DE3) cells at 30˚C. ..................................... 171
Figure 7.5: Proportion of the plasmid-bearing HMS174 (DE3) cells at 30˚C. ............. 172
Figure 7.6: Growth profiles of HMS174 (DE3) cells at 22˚C. .................................... 173
Figure 7.7: Proportion of the plasmid-bearing HMS174 (DE3) cells at 22˚C. .......... 174
Figure 7.8: Growth profiles of HMS174 (DE3) cells cultured in bioreactor/flask at 22˚C .... 175
Figure 7.9: Proportion of the plasmid-bearing HMS174 (DE3) cells cultured in bioreactor/flask at 22˚C ...................................................... 176
Figure 7.10: Relative gene expressions in HMS174 (DE3) (pBCS + pCMP), 3 and 18 hr after induction by 0.025 mM IPTG at 22°C. ............ 178
Figure 7.11: Congo binding analysis of E. coli HMS174 (DE3) (1) and E. coli HMS174 (DE3) (pBCS + pCMP) (2) on agar plates. ...................... 179
Figure 7.12: Congo binding analysis of E. coli HMS174 (DE3) (1) and E. coli HMS174 (DE3) (pBCS + pCMP) (2) in liquid culture 18 hours after induction ................................................................. 180
Figure 7.13: SDS-PAGE analysis of periplasmic (A) and insoluble (B) proteins extracted from E. coli HMS174 (DE3) (pBCS + pCMP). ................. 182
Figure 7.14: Cultures of E. coli HMS174 (DE3) (1), E. coli HMS174 (DE3) (pBCS + pCMP) (2), and A. xylinum (3). ................................. 183
Figure 7.15: Scanning electron micrographs showing the microstructure of BC fibers produced by A. xylinum (1) and E. coli HMS174 (DE3) (pBCS + pCMP) (2,3,4,5) .................. 184
Figure 7.16: ATR-FTIR spectra of biosynthesised BC by HMS174 (DE3) (pBCS + pCMP) .... 187
Figure 7.17: ATR-FTIR spectra of biosynthesised BC by A. xylinum [110] .................. 187
Figure 8.1: Growth profiles of C41 (DE3) cells at 22˚C. ........................................ 201
Figure 8.2: Proportion of the plasmid-bearing C41 (DE3) cells at 22˚C. ................. 202
Figure 8.3: Growth profiles of C41 (DE3) cells at 30˚C. ....................................................... 203
Figure 8.4: Proportion of the plasmid-bearing C41 (DE3) cells at 30˚C. ............................. 204
Figure 8.5: Growth profiles of C41 (DE3) cells at 37˚C. ....................................................... 205
Figure 8.6: Proportion of the plasmid-bearing C41 (DE3) cells at 37˚C. ............................. 206
Figure 8.7: Relative gene expressions of C41 (DE3) (pBCS + pCMP), 3, 6 and 18 hr after induction by 0.05 mM IPTG at 30˚C. ................................................................. 207
Figure 8.8: Relative gene expressions of C41 (DE3) (pBCS + pCMP) 3, 6 and 18 hr after induction by 1.0 mM IPTG at 30˚C. ................................................................. 208
Figure 8.9: Cultures of native strain (control) (1) and genetically modified strain (experiment) (2). .................................................................................................................. 210
Figure 8.10: Scanning electron micrographs showing the microstructure of BC produced by E. coli C41 (DE3) (pBCS + pCMP) after 24 hr fermentation (1,2,3,4), E. coli O157:H7 (5), E. coli PHL644 (6)[245, 246]. .................................................................................................... 211
Figure 8.11: ATR-FTIR spectra of biosynthesised BC by C41 (DE3) (pBCS + pCMP). ............ 213
Figure 9.1: The glucose (1) and acetate (2) metabolism in E. coli [266, 267]. ..................... 220
Figure 9.2: Activity of BcsAB complex in a pH range from 4.5 to 9.5 [271]. ......................... 222
Figure 9.3: Growth profiles of C41 (DE3) cells in LB media supplemented by two different carbon sources. ........................................................................................................ 227
Figure 9.4: Consumption of carbon sources C41 (DE3) cells in LB media supplemented by two different carbon sources. ................................................................. 228
Figure 9.5: Acetate production of C41 (DE3) cells in LB media supplemented by two different carbon sources. .................................................................................. 229
Figure 9.6: pH profile of C41 (DE3) cells in LB media supplemented by different carbon sources. ........................................................................................................ 230
Figure 9.7: Framework of the model development. ................................................................. 234
Figure 9.8: Comparison of growth profile between model simulation and experimental data in batch cultures. ........................................................................................................ 240
Figure 9.9: Comparison of glucose and acetate concentrations between model simulations and experimental data for Experiment 1. ................................................................. 240
Figure 9.10: Comparison of growth profile and glucose concentration between model simulations and experimental data for Experiment 2 ................................................. 242
Figure 9.11: Comparison of growth profile and glucose concentration between model simulations and experimental data in fed-batch cultivation in flasks. .................................................. 243
Figure 9.12: pH profile of fed-batch cultures. ....................................................................... 244
Figure 9.13: Comparison of growth profile and glucose concentration between model simulations and experimental data in fed-batch cultivation in bioreactor. .................................................. 246
Figure 9.14: Microscope images at different magnifications of the biofilm produced in this study (A-N). Microscope images of biofilm produced by Salmonella Typhimurium MAE52 (O and P). AFM images of biofilm produced by Salmonella Typhimurium MAE52 (R and S)[289]. ........................................................................................................ 247
Figure 9.15: Direct observations of the biofilm produced in this study: length (a-b), liquid culture (c) .................................................................................................................. 250
LIST OF TABLES

Table 2.1: Comparison of features of bacterial cellulose and plant-derived cellulose. 29
Table 3.1: The BC productivities of various cellulose-producing bacteria. Adapted from Chawla et al. [99]. 49
Table 4.1: E. coli strains and plasmid vectors that used in this study. 53
Table 4.2: Composition of LB Media. 54
Table 4.3: Composition of LB Agar Media. 55
Table 4.4: Composition of HS Media. 55
Table 4.5: Composition of HS Agar Media. 56
Table 4.6: SDS-PAGE Fixing Solution. 56
Table 4.7: SDS-PAGE Washing Solution. 57
Table 4.8: SDS-PAGE Staining Solution. 57
Table 4.9: Kits used in this study. 58
Table 4.10: Reaction mixture for PCR. 73
Table 4.11: Reaction steps for PCR. 73
Table 4.12: Primers designed for Real Time q-PCR. 74
Table 4.13: Reaction steps for PCR. 77
Table 4.14: Reaction steps for PCR. 79
Table 4.15: Run method for q-PCR. 80
Table 4.16: Reaction components for restriction digest test. 81
Table 4.17: Composition of reaction mixture for glucose determination assay. 84
Table 4.18: Composition of reaction mixture for glycerol determination assay. 86
Table 4.19: Composition of reaction mixture for acetate determination assay. 87
Table 5.1: Important factors for a successful recombinant expression. 91
Table 5.2: Widely used origin of replications. 95
Table 5.3: Origin of replications and copy numbers of Duet vectors. 100
Table 5.4: Sequence landmarks of recombinant expression vectors. 107
Table 6.1: Alterations in E. coli to plasmid-imposed metabolic burden at cultural and cellular level. 124
Table 6.2: E. coli strains and plasmid vectors used in this chapter. 130
Table 6.3: Control and experimental groups used in Chapter 6. 131
Table 6.4: Experimental conditions used in Chapter 6. 132
Table 7.1: E. coli strains and plasmid vectors used in Chapter 7. 167
Table 7.2: Control and experimental groups used in Chapter 7. 168
Table 7.3: Experimental conditions used in Chapter 7. 168
Table 7.4: Main absorption bands and their assignments for the ATR-FTIR spectra of bacterial cellulose. 186
Table 8.1: Comparison of C41 (DE3) and C43 (DE3) cells with the parental strain BL21 (DE3) in transformation and expression of heterologous proteins. 195
Table 8.2: E. coli strains and plasmid vectors used in Chapter 8. 198
Table 8.3: Control and experimental groups used in Chapter 8. 200
Table 8.4: Experimental conditions used in Chapter 8. 200
Table 9.1: pH ranges for cell growth, plasmid production and protein production. 221
Table 9.2: Comparison of growth characteristics of C41 (DE3) cells in LB media supplemented by different carbon sources.

Table 9.3: Summary of the parameters used for the development of the model.

Table 9.4: Initial conditions for Experiment 1.

Table 9.5: Initial conditions for Experiment 2.

Table 9.6: Initial conditions for fed-batch cultivation in the bioreactor.

Table 9.7: Comparison of growth characteristics of C41 (DE3) cells in fed-batch cultures.

Table 10.1: Summary of maximum specific growth rates, dcw and product concentrations under various conditions studied in this work.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>acs</td>
<td><em>Acetobacter</em> cellulose synthesis</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>ATR-FTIR</td>
<td>Attenuated total Reflection - Fourier Transform Infrared</td>
</tr>
<tr>
<td>BC</td>
<td>Bacterial Cellulose</td>
</tr>
<tr>
<td>bcs</td>
<td>Bacterial Cellulose Synthesis</td>
</tr>
<tr>
<td>bp</td>
<td>Base Pair</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>C₆H₁₀O₅</td>
<td>Cellulose</td>
</tr>
<tr>
<td>CCP</td>
<td>Cellulose Complementary Protein</td>
</tr>
<tr>
<td>c-di-GMP</td>
<td>Cyclic diguanylic acid</td>
</tr>
<tr>
<td>CTAB</td>
<td>Cetrimonium Bromide</td>
</tr>
<tr>
<td>DA</td>
<td>Dalton</td>
</tr>
<tr>
<td>DAP</td>
<td>Dihydroxyacetone Phosphate</td>
</tr>
<tr>
<td>dgc</td>
<td>Diguanylate Cyclase</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>DNAse</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTPs</td>
<td>Deoxynucleotides</td>
</tr>
<tr>
<td>DO</td>
<td>Dissolved Oxygen</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>ESPA</td>
<td>Sodium N-ethyl-N-(3-sulfopropyl)-m-anisidine</td>
</tr>
<tr>
<td>GK</td>
<td>Glycerol Kinase</td>
</tr>
<tr>
<td>Glc-1-P</td>
<td>Glucose-1-phosphate</td>
</tr>
<tr>
<td>Glc-6-P</td>
<td>Glucose-6-phosphate</td>
</tr>
<tr>
<td>GM</td>
<td>Genetically Modified</td>
</tr>
<tr>
<td>GPa</td>
<td>Giga Pascal</td>
</tr>
<tr>
<td>GPO</td>
<td>Glycerol Phosphate Oxidase</td>
</tr>
<tr>
<td>G6PDH</td>
<td>Glucose-6-phosphate Dehydrogenase</td>
</tr>
<tr>
<td>HCDC</td>
<td>High cell density cultivation</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>His-tag</td>
<td>Polyhistidine-tag</td>
</tr>
<tr>
<td>HS Medium</td>
<td>Hestrin-Shramm Medium</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>kbp</td>
<td>Kilo Base Pair</td>
</tr>
<tr>
<td>kDA</td>
<td>Kilo Dalton</td>
</tr>
<tr>
<td>KH₂PO₄</td>
<td>Potassium Phosphate Monobasic</td>
</tr>
<tr>
<td>KU</td>
<td>Kilo Unit</td>
</tr>
<tr>
<td>LB Medium</td>
<td>Lysogeny Broth Medium</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>Magnesium Sulfate Heptahydrate</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-morpholinopropane-1-sulfonic acid</td>
</tr>
<tr>
<td>MPa</td>
<td>Mega Pascal</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>Disodium Phosphate</td>
</tr>
<tr>
<td>NAD</td>
<td>Nicotinamide Adenine Dinucleotide</td>
</tr>
<tr>
<td>NADP</td>
<td>Nicotinamide Adenine Dinucleotide Phosphate</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>(NH₄)₂SO₄</td>
<td>Ammonium Sulfate</td>
</tr>
<tr>
<td>OD₆₀₀</td>
<td>Optical Density at 600 nm</td>
</tr>
<tr>
<td>-OH</td>
<td>Hydroxyl Group</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>q-PCR</td>
<td>Quantitative Polymerase Chain Reaction</td>
</tr>
<tr>
<td>POD</td>
<td>Peroxidase</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>RNAse</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>rpm</td>
<td>Rotation Per Minute</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse Transcription</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>SEM</td>
<td>Scanning Electron Microscopy</td>
</tr>
<tr>
<td>SOC Medium</td>
<td>Super Optimal broth with Catabolite Repression</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris base, Acetic Acid and EDTA</td>
</tr>
<tr>
<td>TCA</td>
<td>Trichloroacetic Acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>U</td>
<td>Unit</td>
</tr>
<tr>
<td>UDPGlc</td>
<td>Uridine Diphosphate Glucose</td>
</tr>
<tr>
<td>UGPase</td>
<td>Uridine Diphosphate Glucose Pyrophosphorylase</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>4-AAP</td>
<td>4-aminoantipyrine</td>
</tr>
<tr>
<td>UGPase</td>
<td>Uridine Diphosphate Glucose Pyrophosphorylase</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>4-AAP</td>
<td>4-aminoantipyrine</td>
</tr>
</tbody>
</table>
1. GENERAL INTRODUCTION
1.1. Overview

Cellulose is the most abundant polymer found in nature as a structural component, often bound to other polymers (pectin, lignin, arabinan, etc.) in the cell wall of plants, algae and also of some lower animals. It can be also produced by various species of bacteria, such as those of the genera *Gluconacetobacter* (formerly *Acetobacter*), *Agrobacterium*, *Aerobacter*, *Achromobacter*, *Azotobacter*, *Rhizobium*, *Sarcina*, and *Salmonella*.

Bacterial cellulose (BC) is chemically identical to plant cellulose but macromolecular structure and properties are remarkably different. BC exhibits unique physical, chemical and mechanical properties such as high crystallinity, large surface area and elasticity as a result of its nanofibrillar network structure contrary to plant cellulose. It also exhibits higher tensile strength and water holding capacity compared to plant cellulose. In addition to all these properties, another remarkable feature of BC is its chemical purity, which distinguishes it from plant cellulose. Since plant cellulose is associated with hemicelluloses and lignin, an inherently difficult purification process is required for the removal of these substances.

Because of the unique properties, resulting from the ultrafine reticulated structure, there is a tremendous amount of attention for BC. It has a great potential for a wide range of applications in paper, textile, and food industries, and more importantly for biomaterials engineering, medicine, membranes and polymer nanocomposites. This has been confirmed by recent increase of published papers on bacterial cellulose.
Figure 1.1: Summary of the annual number of publications related to bacterial cellulose from the year 2000 to 2015 indexed within the Web of Science®. The keywords for the search included following terms: bacterial cellulose, microbial cellulose.

However, wider application of this polysaccharide is dependent on the scale of production and its cost. Relatively high cost of the production of BC limits the use of BC in industry. Therefore, it is important to develop both economical and high quality BC production methods for the promotion of its use. Recently, intensive studies have focused on strain improvement and bioprocess engineering to solve this problem.

1.2. Motivation and Novelty

Acetobacter xylinum is the most efficient producer of BC. So far, most of the studies have been focused on A. xylinum, however, it has many weaknesses such as slow growth compared to most other bacteria such as E. coli, low productivity and its susceptibility to culture conditions. These factors limit the cost-effective production of BC. Therefore, it is crucial to obtain valuable strains for industrial use. Genetic modification is a powerful tool that could be utilised in obtaining valuable strains for the improvement of BC biosynthesis. E. coli has a significantly fast growth rate with a 20 minutes doubling time, whereas the doubling time of A. xylinum is 4
hours in agitated cultures. Genetic modification of \textit{E. coli} by recombinant DNA technology could be a very promising line of advance towards increasing BC production. Although it has been mentioned in literature that \textit{E. coli} could have a great potential over \textit{A. xylinum} for production of BC, this has not been studied because of the complexity of BC biosynthesis pathway. In an effort to meet this demand, the novel study reported in this thesis aimed and achieved to develop cellulose producing \textit{E. coli} strains by recombinant DNA technology. Optimisation of the bioprocess conditions for an efficient and stable recombinant expression system was also studied to improve BC productivity.

1.3. Roadmap of the thesis

The present thesis is divided into 10 chapters, as follow:

- **Chapter 1** introduces the research context, motivation and novelty of the study and the roadmap of this thesis.

- **Chapter 2** provides a background to support this study with relevant research evidence and arguments based on literature.

- **Chapter 3** describes the aims and the objectives of this study by identification of the research challenges and proposed solution.

- **Chapter 4** provides a background and introduction to methods and describes materials used in this study.

- **Chapter 5** describes the design of plasmid vectors based on important factors for plasmid design (strain-plasmid compatibility, plasmid size, copy number, selective markers, promoter, codon optimisation, GC content). Cloning and transformation strategies are explained. Full sequences and properties of the plasmids, restriction digest test results, GC contents are provided.

- In **Chapter 6**, bacterial cellulose production by genetically modified (GM) \textit{E. coli} Strain 1 (\textit{E. coli} BL21 (DE3)) was investigated by optimisation of bioprocess
parameters (temperature, IPTG concentration, induction OD$_{600}$). Growth profile and plasmid stability of the cells are presented. Expression of the target genes and proteins are reported. Congo red binding assay results are provided.

- In **Chapter 7**, the successful BC production by genetically modified *E. coli* Strain 2 (*E. coli* HMS 176 (DE3)) is reported. Optimisation of bioprocess parameters (temperature, IPTG concentration) by monitoring growth rate and plasmid stability is presented. Expression levels of the target genes and plasmid stabilities are determined. SEM and ATR-FTIR and Congo red binding assay results are provided.

- In **Chapter 8**, BC production by genetically modified *E. coli* Strain 3 (*E. coli* C41 (DE3)) is presented. Optimisation of bioprocess parameters (temperature, IPTG concentration) by monitoring growth rate and plasmid stability is presented. Expression levels of the target genes and plasmid stabilities are determined. Characterisation of the product by SEM and ATR-FTIR are shown.

- In **Chapter 9**, BC production by Genetically Modified *E. coli* C41 (DE3) in flasks and stirred-tank bioreactor is reported under fed-batch conditions. Fermentation of the GM cells in the presence of various concentrations of glucose and glucose are presented. A mathematical model, which is created to determine feeding strategy, is provided.

- General conclusions are drawn and future work is discussed in **Chapter 10**.
2. GENERAL BACKGROUND
2.1. Cellulose Types

Cellulose is a significantly important biopolymer for a wide range of industrial applications such as medical, food, paper, and cosmetics. In addition being the main component of the plant cell wall, it can be produced by various species of bacteria.

2.1.1. Plant Cellulose

Cellulose was first described in 1838 by the French chemist Anselme Payen as a resistant fibrous material that obtained after treatment of various plant tissues with acids and ammonia, and following extraction with water and alcohol [1]. He obtained the molecular formula of this solid as \( \text{C}_6\text{H}_{10}\text{O}_5 \) by elemental analysis, and observed the isomerism with starch. The term “cellulose” for plant constituent was first used in 1839 in a report of the French academy that published the study of Payen [2].

Throughout the plant kingdom, the cell wall is very strong as cellulose fibrils are the primary architectural elements [3]. The plant cell wall is a semipermeable tissue that compounded from layered deposits of cellulose fibrils in which are embedded more amorphous polymers, including neutral and acidic polysaccharides, glycoproteins, and waxy aromatic substances [4]. Based on electron-microscopic images of plant cell wall [5], cellulose is a strong supporting material since it has a high tensile strength comparable to that of steel. Cellulose fibrils are highly insoluble and inelastic. As a result of properties, cellulose is exhibits a unique combination of chemical resilience and mechanical support and flexibility to the tissues, such as paper, lumber or cotton textiles. [6] Cellulose has been used for about 150 years as a chemical raw material in different fields. About 33% of all plant matter is cellulose (the cellulose content of cotton is 90% and that of wood is 40–50% [7].
2.1.2. Bacterial Cellulose

Bacterial cellulose produced by *Acetobacter xylinum* was described initially in 1886 by A. J. Brown as an extracellular gelatinous mass formed during the vinegar fermentation [9]. However, it became a significant research topic in the second half of the 20th century [10]. From then onwards, bacterial cellulose produced by *A. xylinum* is clearly identified with an extensive research in terms of structure, formation and fermentation conditions. The most efficient bacterial cellulose producer is gram-negative *A. xylinum* with the ability of converting 50% per cent of a wide range of carbon and nitrogen sources to extracellular cellulose, which is easy to isolate as a fiber material [11]. Consequently, *A. xylinum* has become most commonly studied source of bacterial cellulose and has been applied as model microorganism for basic and applied studies on cellulose. Gram-negative species such as *Agrobacterium, Achromobacter, Aerobacter, Enterobacter, Sarcina, Rhizobium, Pseudomonas, Salmonella and Alcaligenes* have also been found to produce cellulose [12-17]. There are also reports on some Gram-positive species such as *G. hansenii* that have ability to synthesise cellulose as well [18].
The majority of bacteria produce extracellular polysaccharides, which form envelope-like structure around cells [19]. Similarly, cellulose-producing bacteria are embedded in the cellulose network, which supports the population at the liquid-air interface [20]. Cellulose layer helps nutrient supply for embedded bacteria, as their concentration in the polymer matrix is significantly enhanced due to highly adsorptive structure. [19, 21]. Moreover, cellulose layer protects cellulose-producing cells against critical changes such as pH, water content, and accumulation of toxic substances. It has been reported that cellulose layer protects bacteria from ultraviolet radiation [6].

2.2. Advantages of Bacterial Cellulose over Plant Cellulose

Although plant cellulose has been considered as the most abundant polymer on Earth, bacterial cellulose has outstanding physical, chemical and mechanical properties over plant cellulose due to its molecular structure.

![Glycosidic bond and Hydrogen bonds](image)

**Figure 2.2:** Glycosidic bond between β-D-glucose molecules (A) and hydrogen bonds on the same or the neighbour chain (B).

Molecular structure of cellulose as a carbohydrate polymer consists of repeating β-D-glucose molecules. These molecules are covalently linked through acetal functions between the equatorial -OH group of C4 and the C1 carbon atom.
(β-1,4-glucan bonds). As a result, cellulose is a linear-chain polymer with a large number of hydroxyl groups. The polar -OH groups form many hydrogen bonds with oxygen atoms on the same or on a neighbour chain, as presented in Figure 2.2. The hydrogen bonds between and within cellulose chains constitute stable crystalline regions and give the structure more stability and strength. Two forms of cellulose are produced by Acetobacter xylinum: (i) cellulose I, the ribbon-like polymer, and (ii) cellulose II, the thermodynamically more stable amorphous polymer [22]. Assembly of two microfibrillar structures (cellulose I and cellulose II) is significantly different. Despite the identity in the chemical structure of bacterial and plant cellulose, the bacterial cellulose differs from plant cellulose in terms of microfibrillar structure. Although both bacterial and plant cellulose have identical molecular formula, microfibrillar structure of bacterial cellulose leads to most of its properties.

Bacterial cellulose has excellent mechanical properties, over plant-derived cellulose, owing to crystalline nano- and microfibril structure such as high purity, high degree of polymerisation and crystallinity index, high tensile strength, and water holding capacity, as presented in Table 2.1. [23]. These unique properties allow bacterial cellulose to be used in a wide range of industrial applications such as electronics, medical products, paper, food, cosmetics and medicine. High modulus of elasticity in combination with a large internal loss factor gives the bacterial cellulose an important role as a more suitable raw material for the production of high fidelity acoustic speakers (Sony Corp.). Fibrils of bacterial cellulose are approximately 100 times thinner than that of plant cellulose, resulting in a notably porous material. This enables the transfer of antibiotics or other medicines into the wound and serves as an effective physical barrier against any external infection at the same time. Therefore it has extensive use in wound healing[24].
Table 2.1: Comparison of features of bacterial cellulose and plant-derived cellulose.

<table>
<thead>
<tr>
<th>Features</th>
<th>Plant cellulose</th>
<th>Bacterial Cellulose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purity</td>
<td>exists in a mixture with lignin, pectin and hemicellulose.</td>
<td>highly pure</td>
</tr>
<tr>
<td>Purification Process</td>
<td>environmentally-unfriendly</td>
<td>environmentally friendly</td>
</tr>
<tr>
<td></td>
<td>spend high energy</td>
<td>simple purification</td>
</tr>
<tr>
<td>Crystallinity Index</td>
<td>43% to 65%</td>
<td>70% to 90%</td>
</tr>
<tr>
<td>Degree of Polymerisation</td>
<td>up to 14000</td>
<td>up to 20000</td>
</tr>
<tr>
<td>Tensile Strength</td>
<td>530 to 675 MPa</td>
<td>78 to 155 GPa</td>
</tr>
<tr>
<td>Water Holding Capacity</td>
<td>up to 8%</td>
<td>up to 99%</td>
</tr>
</tbody>
</table>

Bacterial cellulose is characterized by its high purity, while plant cellulose exits with substances like hemicelluloses, lignin, or pectin. It has an extremely high water content of 90% or more. Purification process for plant cellulose has mechanical and chemical separation steps including logging, debarking, chipping, mechanical pulping, screening, chemical pulping and bleaching, which require a high energy and the whole purification process itself is environmentally unfriendly (Figure 2.3). The microbial cellulose obtained after fermentation is highly pure; but it contains some impurities such as cells and/or the medium components. However the purification process is extremely simple compared to purification of plant cellulose. Several types of widely used purification process of bacterial cellulose in the culture medium include the treatment with alkali (sodium hydroxide or potassium hydroxide), organic acids like acetic acid or repeated washing of the mixtures with the reverse osmosis water [11, 25].

2.3. Metabolic Pathway of Cellulose-Producing Bacteria

Cellulose-producing bacteria can operate the pentose-phosphate cycle or the Krebs cycle, depending on the physiological state of the cell coupled with gluconeogenesis [6]. The oxidation of carbohydrates is achieved by the pentose-phosphate cycle. Krebs cycle involves the oxidation of acetate-derived
carbohydrates, fat and proteins, such as oxalosuccinate and α-ketoglutarate. Since *A. xylinum* lacks phosphofructose kinase, it is not able to metabolise glucose anaerobically, which is essential for glycolysis [26]. The biosynthesis of cellulose by *A. xylinum* has been widely studied [3, 27-32]. It is a multi-step reaction involving individual enzymes, catalytic complexes and regulatory proteins. If glucose is used as carbon source, biosynthesis pathway constitutes of four key enzymatic steps (Figure 2.4); they are: (i) phosphorylation of glucose by glucokinase (ii) isomerization of glucose-6-phosphate (Glc-6-P) to glucose-1-phosphate (Glc-1-P) by phosphoglucomutase, (iii) synthesis of UDP-glucose (UDPGlc) by UDPG-pyrophosphorylase (UGPase) and (iv) cellulose synthase reaction. UDPGlc is a common molecule in many organisms, which is the direct cellulose precursor. UGPase is approximately 100 times more active in cellulose producers than that of non-cellulose producing bacteria. As a result of this, it is thought to play an important role in cellulose synthesis. If disaccharides, such as sucrose and maltose, are used as carbon source for cellulose-producing bacteria, the first step of biosynthesis of bacterial cellulose is hydrolysis of disaccharides into...
monosaccharides, such as glucose and fructose. Although pathways of UDPGlc are relatively well known, the mechanisms of glucose polymerisation into long and unbranched cellulose chains are still elusive to scientists [32].

Cyclic diguanylic acid (c-di-GMP) has also an important role in the synthesis of BC as an allosteric activator for the cellulose synthase. In the absence of c-di-GMP, cellulose synthase remains in an inactive state or exhibits low enzyme activity [6, 33]. Another important protein for the pathway is a membrane protein: c-di-GMP binding protein. It is found structurally associated with the cellulose synthase; 90% of the cellular c-di-GMP is reversibly bound by the c-di-GMP binding protein. The balance between bound and free c-di-GMP is regulated by the intracellular potassium concentration [33-35].

Except from enzymatic steps, structural assembly of cellulose fibres has two intermediary steps: (i) the formation of 1,4-β-glucan chains and (ii) the assembly and crystallisation of cellulose chains. The assembly and crystallisation of cellulose is the rate-limiting step of the process [36]. BC formation takes place in between the outer and cytoplasm membranes of the cell (Figure 2.4) [37]. First, cellulose molecules are synthesised inside the bacteria. These molecules are then spun through cellulose export components to form of protofibrils. These fibrils are approximately 2-4 nm in diameter and a ribbon shaped microfibril, which is approximately 80 nm in diameter, is assembled from these protofibrils [38].

It has been known that, the biosynthesis of cellulose is catalysed by cellulose synthase, which polymerises the glucose units into the 1,4-β-glucan chains, in the last step of enzymatic multi-step reaction [36]. However, the polymerisation mechanism of glucose monomers into glucan chains is not yet well understood. One hypothesis is that the glucose residues are attached onto the non-reducing end of the polysaccharide, which takes place in the extracytoplasmic space. Another plausible hypothesis that the process involves a lipid intermediate, which cellulose chains run through the centre of transmembrane, during the polymerisation of 1,4-β-glucan [39].
2.4. Genetic Pathway of Acetobacter

Recently, studies have focused on application of the genetic modifications to bacteria in order to raise bacterial cellulose production. As a result, molecular basis of the bacterial cellulose production has become more of interest. *Acetobacter* cellulose synthesis operon (*acsABCD*) and bacterial cellulose synthesis operon (*bcsABCD*) are two homologous operons that exist in *A. xylinum* ATCC 53582 and 1306-3, respectively. Both operons encode the essential proteins for cellulose synthesis [32, 41]. Cellulose synthase, which is involved in synthesis of cellulose from UDP-glucose, is encoded by three (*acsAB, acsC, and acsD*) or four (*bcsA, bcsB, bcsC, and bcsD*) subunits. The first gene of the *bcsABCD* operon, *bcsA*, encodes the catalytic subunit of cellulose synthase and binds to UDPglc. The second gene, *bcsB*, encodes the regulatory subunit of cellulose synthase that binds to c-di-GMP. It also plays an important role as second messenger and activates the cellulose synthesis process [41]. *acsA* and *acsB* encode a single polypeptide that has both substrate binding and activator-binding regions. However, the functions of *acsC/bcsC* and *acsD/bcsD* have not been completely clarified yet. It has been suggested that
acsC/bcsC is responsible for the formation of pores to secrete cellulose since acsC/bcsC encodes proteins that are similar to the proteins involved in membrane channels or pore formation [32]. Deactivation of acsA, acsB and acsC blocks the synthesis of BC completely, which shows that these genes are essential for the production. On the other hand, the deactivation of acsD decreases cellulose production by 40% [32, 36]. This suggests that acsD controls the crystallisation of cellulose into nanofibrils [42].

The upstream region of the acs/bcs operon has two genes; cmcax and ccpAx, respectively (Figure 2.5). cmcax encodes endo-β-1,4-glucanase enzyme, which has cellulose hydrolysing activity. Although the function of Cmcax in cellulose biosynthesis have not been identified, studies unexpectedly showed that endo-β-1,4-glucanase enzyme enhances cellulose synthesis [43-45]. Based on an electron microscopic analysis, Kawano et al. suggested that Cmcax from A. xylinum could influence in cellulose ribbon assembly. Their data revealed that the cellulose ribbons secreted from the Cmcax overproducing strain were dispersed compared with those from the wild type strain. Kawano et al. suggested a regulation mechanism of Cmcax expression in a non-cellulose producing mutant of A. xylinum [46]. The other protein in the upstream region of acs operon is CcpAx, which is essential for the production with a complementing function. It also enhances the production of BC. However the nature of this function still remains to be elucidated [47]. It has been suggested that the functions of CcpAx could be related to transport of cellulose chains from the sites of polymerisation within the cell membrane to the extracellular environment and the crystallization process of cellulose fibrils [17]. A sequence of IS1031 element upstream of the start of the transcription of this operon was inserted to investigate the effect on BC biosynthesis [48]. As a result, cellulose deficiency was observed in the mutant strain. This shows that the upstream region of the operon is important for the synthesis of BC. The downstream region carries the gene bg1xA that encodes β-glucosidase, which has activity in hydrolysis on more than three β-1,4-glucose units. It was shown that the disruption of the bg1xA gene causes a decrease in BC.
production [46]. A schematic representation of the genetic pathway of the cellulose synthase related genes in *A. xylinum* ATCC 53582 is presented in Figure 2.5.

![Schematic diagram showing the genetic pathway of *A. xylinum* ATCC 53582.](image)

**Figure 2.5**: A schematic diagram showing the genetic pathway of *A. xylinum* ATCC 53582. Regions 1, 2 and 3 represent the upstream of cellulose synthase operon, cellulose synthase operon, and the downstream cellulose synthase operon, respectively (Adapted from Lee et al. [40]).

### 2.5. Fermentative Production of BC

BC productivity is mainly affected by culture conditions, such as the composition of the culture medium, pH, temperature, dissolved oxygen content and the culturing strategy (static or agitated fermentations). The optimal design of fermentation conditions is crucial for the growth of cellulose producing bacteria to stimulate the formation and production of BC.

#### 2.5.1. Composition of Culture Media

One of the most important factors affecting the BC yield is the type of carbon source used for the fermentation of cellulose producing bacteria. Glucose and sucrose are the most commonly used carbon sources for BC production. Other types of carbohydrates, such as fructose, maltose, xylose, starch and alcohols, such as glycerol, mannitol and arabinol have also been studied [18, 49, 50].
The optimum culture medium for the production of BC was studied by Bae and Shoda [51]. Box-Behnken design was used to optimise the concentration of various components within the culture medium. According to the results, a BC concentration of 14 g L\(^{-1}\) can be obtained after 72 h fermentation time when using a culture medium containing 4.99 wt.-% of fructose, 2.85 wt.-% corn steep liquor, which is a viscous liquid by-product of corn wet milling, rich in amino acids, vitamins and other minerals, 28.33 wt.-% dissolved oxygen content and 0.38 wt.-% agar. Further study done by the same group showed that when the carbon source is changed to H\(_2\)SO\(_4\) treated molasses, which is a viscous by-product from sugarcane refining, the BC yield produced by \emph{A. xylinum} BPR2001 increased by 76% compared to neat molasses [52]. Jonas and Farah used numerous mono-, di-, polysaccharides, alcohols (ethanol, glycerol, ethylene and glycol), organic acids (citrate, succinate, gluconate), and other compounds (glucono-lactone, O-methyl-glucose) to compare the effect of carbon source on the BC yield. They claimed that the preferred carbon sources for BC production were D-arabitol and D-mannitol, which resulted in a 6.2-fold and 3.8-fold greater BC yield, respectively, compared to glucose [21].

Although glucose is the most widely used carbon source for the cultivation of cellulose-producing bacteria, the formation of gluconic acid can be problematic. Gluconic acid is formed as a by-product of metabolising glucose; consequently, accumulation of gluconic acid decreases the pH of the culture and also the production of cellulose. Therefore, initial glucose concentration is an important parameter for BC yield. Masaoka \textit{et al.} investigated the BC yield of \emph{A. xylinum} IFO 13693 at various glucose concentrations of 6, 12, 24 and 48 g L\(^{-1}\), respectively. It was reported that the BC yield decreases with increasing initial glucose concentration in the culture medium [53]. The amount of consumed glucose equals to the total BC and gluconic acid production in the culture. This suggested that if glucose is not consumed for cellulose synthesis, it can be metabolized via gluconic acid to other substances.

When the medium was supplied by maltose as a carbon source, the BC yield was 10 times lower than that of a culture medium containing glucose as the carbon
source [53]. Matsuoka et al. have also showed that when lactate is used in the culture medium, the growth of *A. xylinum* ssp. sucrofermentous BPR2001 under agitated conditions increases [50]. Also, BC yield increased by approximately 4-5 times. It was suggested that lactate acts not only as an accelerator to drive the tricarboxylic acid (TCA) cycle, but also an energy source for *A. xylinum* ssp. sucrofermentous BPR2001. These two could be the reasons behind the more rapid cell growth and higher BC yield.

The culture medium supplied by ethanol is also advantageous for the production of BC. Ethanol has ability to supress the spontaneous mutation of cellulose producing bacteria into cellulose non-producing mutants [25]. This usually appears under agitated culture conditions. Ethanol also used as additional carbon source for *G. hansenii* [54]. The results showed that the BC yield by *G. hansenii* increased from 1.30 to 2.31 g L$^{-1}$ with of 1 vol.-% ethanol supplement. *Acetobacter* sp. A9 strain was also used by Son et al. to investigate the effect of ethanol on the BC production. The culture medium supplemented with 1.4 vol.-% ethanol showed 400% increase in the BC concentration (15.2 g L$^{-1}$) compared to culture medium, which did not contain ethanol. This significant increase in BC concentration can be attributed to the aforementioned benefits of ethanol.

A nitrogen source is also an important component of culture medium as it provides crucial amino acids to the bacteria. The most widely used nitrogen sources are yeast extract and peptone, which are the basic components of the model medium for BC production developed by Hestrin and Schramm [10]. However, it has been shown that the most recommended nitrogen source for agitated cultures is corn steep liquor [50]. When corn steep liquor was provided it was added in low concentrations (0.15 vol.-%) to the medium containing 4 % (wt./vol.) of fructose, the results showed that BC production was stimulated. Corn steep liquor contains lactate, which is absent in other nitrogen sources, is the main reason for the stimulation of BC production. Ramana et al. studied the effect of the presence of different nitrogen sources [55]. When casein hydrolysate was provided to the culture medium, a BC concentration of 5 g L$^{-1}$ was obtained. On the other hand,
peptone as the nitrogen source yielded only 4.8 g L\(^{-1}\) of BC. Studies on the addition of extra nitrogen showed that the biomass and BC production were enhanced. The influence of vitamins, such as pyridoxine, nicotinic acid, p-aminobenzoic acid and biotin, on BC production was also studied. The results showed that these vitamins were the most stimulating vitamins for BC production. However, pantothenate and riboflavin have a negative effect on the BC productivity [50, 56].

### 2.5.2. Bioprocess Parameters

pH, temperature and dissolved oxygen content are the main environmental parameters that affect growth and metabolism of cellulose producing bacteria. When any of these parameters are changed, microorganisms rapidly respond in terms of induction and repression of protein synthesis and changes in cell morphology.

### 2.5.3. Influence of pH on BC Production

The optimum pH for the growth of bacteria and production of BC is specific for every particular strain of bacteria. Although it is usually in the range of 4 to 7, the highest BC production obtained at pH 6.5. [54, 57]. However, according to the industrial producers (Biofill and Gengiflex) of BC membranes for biomedical applications, was worked at low pH of between 4.0-4.5 since under this condition it is possible to avoid contamination of the medium during BC culturing [21]. Maintaining the pH of the culture medium for the maximum yield of BC is a bottleneck. As a result of the accumulation of secondary metabolites, such as gluconic, acetic or lactic acids that are produced during the consumption of sugars and nitrogen sources, the pH of the culture medium could decrease as a function of time. Therefore, corn steep liquor can be added into the culture medium as a buffer to maintain the pH of the culture medium [58]. However, corn steep liquor increases the viscosity of the medium, which could cause inhomogeneity in the medium.
2.5.4. Influence of Temperature on BC production

Temperature has effects on growth of cellulose producing bacteria and/or crystal structure of BC. Son et al. studied the influence of temperature (20°C to 40°C) on the yield of BC produced by Acetobacter sp. A9 in Hestrin and Schramm medium [54]. The optimum temperature for BC production was obtained to be 30°C. When they lowered the culture temperature to 25°C, they did observe a significant decrease in the BC yield compared to 30°C. However, it has been showed that increasing the temperature to 35°C reduces the BC yield [41].

The morphology and crystal structure was also affected by culture temperature. Hirai et al. claimed that BC produced by A. xylinum ATCC 23769 in HS medium at 4°C yielded band shaped BC with a cellulose II structure [59]. Meanwhile, BC produced at 28°C by the same strain, yielded cellulose I ribbons. These results supported by the study of Zeng et al.[60]. Cellulose I was produced by A. xylinum BPR2001 in a medium composed of 20 g/L fructose, 3.3 g/L (NH₄)₂SO₄, 20 g/L yeast extract, 1 g/L KH₂PO₄ and 0.122 g/L MgSO₄7H₂O when the culture temperature was maintained between 25°C and 30°C.

2.5.5. Influence of Oxygen on BC production

The dissolved oxygen content in the culture medium is critical for cell metabolism, plasmid stability, and protein expression for bacterial fermentation. It is also extremely important for improving the yield and quality of BC [61]. However, it was reported that high dissolved oxygen content in the medium would result in an increase in the gluconic acid concentration [62]. This would decrease the cell viability, which ultimately reduces the yield of BC. On the other hand, low dissolved oxygen content, inhibits bacteria growth and production of BC. In addition, maximum BC concentration was observed at 10% saturation of dissolved oxygen for batch-fed cultures [63].
2.6. Genetic Modification of Bacteria to Enhance BC Production

The main product of the oxidation of carbon source is ketogluconate for A. xylinum glucose or sucrose is used as carbon source. Cellulose is the byproduct of the oxidation process [53]. Inhibition of the production pathway of ketogluconate and force conversion of the glucose into cellulose could be a solution to increase cellulose production. For this purpose, ketogluconate-negative Acetobacter strains were isolated and this mutant strain was derived by UV-mutagenesis [64]. It was claimed that the BC concentration increased from 1.8 g L\(^{-1}\) (the parent strain) to 3.3 g L\(^{-1}\) after 10 days of cultivation. Meanwhile the consumption of glucose by the mutant strain decreased from 22.6 g L\(^{-1}\) for the parent strain to 7.3 g L\(^{-1}\). The reason behind the decrease in glucose consumption is attributed to the inhibition of the metabolic pathway that converts glucose into ketogluconate.

dgc1 is a gene that plays an important role in activating BC synthesis, which catalyses the synthesis of c-di-GMP. Therefore, genetic modification applied to A. xylinum BPR2001 with the expecting that the disruption of dgc1 should decrease BC production [65]. Contrary to the expectations, the BC production of dgc1-disrupted mutants remained approximately the same compared to the production of the parent strain, in both static and shake flask cultures. However the growth rate of dgc1-disrupted mutants was slower than that of the parental strain. This may be an explanation for why there was no change observed in the overall BC yield of dgc1 disrupted strain under static or agitated conditions. On the other hand, in a stirred tank reactor, dgc1-disrupted mutants was produced BC with the yield of 36% higher than that of the parent strain. This proves that although c-di-GMP synthesis is essential for cellulose synthase activation, disruption of the dgc1 gene may not be fatal for BC synthesis. It was hypothesised the BC production was complemented and enhanced by two other proteins: dgc2 and dgc3, which have similar functions to those of dgc1. On the other hand, according to another study, a decrease reported in BC produced by dgc1-disrupted strain. The contradictory results
obtained between those studies could be because of the short cultivation period used by the latter study to draw a conclusion for the final yield of BC. [65, 66]

*A. xylinum* secretes acetan during the production of BC [67]. Acetan is a viscous water-soluble polysaccharide, produced from UDPGlc, which is also the starting compound to produce cellulose. Therefore, when the production of acetan is inhibited, the concentration of UDPGlc is expected to increase. This enhances the production of BC in return. For this purpose, a non-acetan producing mutant strain (EP1) was derived from *G. xylinus* BPR2001 [68]. Contrary to expectations, the BC production of EP1 decreased compared to the parent strain under agitated conditions. Under static conditions, no significant difference was observed in the yield of BC between EP1 and the parental strain. The culture broth contained large flocs of cells and BC which resulted in heterogeneous suspensions during cultivation of EP1. Because when there is no acetan production, viscosity of the culture medium decreases. Low viscosity increased the likelihood of cell and BC coagulation, which resulted in a decrease in BC production.

DNA fragments carrying cellulose synthesis related genes in the upstream and downstream regions of the bcs operon in *A. xylinum* ATCC23769 and ATCC53582 was sequenced by Kawano *et al.* [69]. The cellulose synthesis related DNA fragments contain endo-β-1,4-glucanase, cellulose complementing protein, cellulose synthase subunits AB, C and D, and β-glucosidase genes. It was reported that after 7 day incubation, the BC yield of ATCC53582 detected 5 times more than the BC yield of ATCC23769. Unexpectedly, ATCC53582 continued to produce BC after the glucose was totally consumed. There are two suggestions to explain this result; either gluconic acid was used as carbon source for the production of BC or; the glyconeogenesis pathway was activated. This draws a conclusion that ATCC23769 consumes its energy for cell growth whilst ATCC53582 uses its energy for BC production.

The proteins in the upstream region of acs operon are CcpAx and CMCax. CcpAx is suggested to be involved in cellulose crystallisation, while CMCax plays role in both cellulose hydrolysis and synthesis of BC [70]. The relationship between the
structure and function of these genes was also investigated [71]. It has been observed a 1.2 fold increase in the yield of BC when CMCax was over-expressed in *A. xylinum*. In addition, when the culture medium was supplemented by CMCax protein, the production of BC increased [46].

Cellulose synthase operon (*acsABCD*) of *G. xylinus* was cloned into unicellular cyanobacteria (*Synechococcus leopoliensis* strain UTCC 100) and the genes were expressed successfully in genetically modified *Synechococcus leopoliensis* [72]. The genetically modified strain produced amorphous cellulose, which is lack of the typical fibrillar structure of BC. A gene sequence encoding a putative pyrroloquinoline quinone glucose dehydrogenase from *G. xylinus* BPR2001 was cloned in order to produce a glutamate dehydrogenase (GDH)-deficient mutant strain of BPR2001 (GD-I) [73]. The GD-I strain does not produce gluconic acid but it produces 4.1 g L\(^{-1}\) of BC aerobically in a medium containing glucose as carbon source. This BC production of GD-I was approximately 2 times higher than that of the wild strain. The yield coefficient values (grams of BC produced per gram of consumed glucose) of strains GD-I and BPR2001 were found to be 0.1 and 0.06, respectively.

Genetic modification is a powerful tool that can be used in the obtaining valuable strains for the production of BC. It could be a potential solution towards increasing BC production to apply genetic modification on *E. coli*.

### 2.7. Recombinant DNA Technology and Molecular Cloning

Biotechnology is a fast developing research field which includes knowledge from a combination of several traditional sciences such as biochemistry, microbiology, chemical engineering and chemistry.

The idea of recombinant DNA was first proposed in the Biochemistry Department at Stanford University by Peter Lobban, a graduate student of Prof. Dale Kaiser. The first studies describing the successful production and intracellular replication of recombinant DNA were published in 1972 and 1973 [74, 75]. The first
licensed drug produced using recombinant DNA technology was human insulin, which was created by Genentech and licensed by Eli Lilly and Company.

In early 21st century, diverse sciences such as genomics, recombinant gene technologies, and applied immunology have been integrated with biotechnology as a result of developments in medicine, food production, materials, and agriculture. Genes can be isolated and amplified to facilitate their use in various species. The most common method of isolation and amplification of a target gene is to clone the sequence of the target gene by inserting it into another DNA molecule that serves as a vehicle such as a vector that can be replicated in cells. Recombinant DNA can be transferred into a very wide range of living organisms. It is possible to insert a plant DNA sequence to a bacterial DNA sequence or a human DNA sequence to a fungal DNA sequence. When these DNAs of different origin are combined, the genetic product is a recombinant DNA molecule. The recombinant DNA molecule is placed in a host cell, which can be either prokaryotic or eukaryotic. When the host cell replicates, the vector carrying target sequence of DNA also replicates. Thus, the target DNA becomes amplified in number, and following its amplification can be purified for further analysis.

Proteins produced by the expression of a recombinant DNA are termed recombinant proteins. The recombinant protein is not necessarily produced when recombinant DNA encoding a protein of interest is introduced into a host organism. Successful expression of foreign proteins requires the use of specialized expression vectors. Also the target sequence requires significant restructuring in order to produce active proteins. Today, a broad range of recombinant expression vectors and compatible host are available in the market.

2.8. Recombinant protein expression in *Escherichia coli* by plasmid vectors

*Escherichia coli* (*E. coli*) is one of the most widely used organisms for the production of recombinant proteins. It is well-characterised microorganism, which allows for a well-established cell factory and it has become the most popular
expression platform. It has fast growth kinetics and high density cultures, which are easily achieved by *E. coli* [76-79]. Rich complex media can be prepared using inexpensive components for the fermentation of *E. coli*. For this reason, the greatest variety of cloning vectors has been developed for use in *E. coli* as the bacterial host. Also, there are many molecular tools and protocols for the high-level production of recombinant proteins, such as a vast catalogue of expression plasmids, a great number of engineered strains and many cultivation strategies.

The most important parameters to choose or design a plasmid are multiple combinations of origin of replications (pUC, pET etc.), promoters (*lac*, T7 etc.), selection markers (ampicillin, kanamycin etc.), multiple cloning sites and affinity tags (His-, GST etc.). These parameters have to be carefully evaluated according to particular need.

At the theoretical level, the steps needed for designing a recombinant protein expression system seem straightforward: i) amplification of gene of interest, cloning the target gene into expression vector, ii) transformation it into the host of choice, iii) induction of the protein expression. However there are many problems that scientists could face. Poor growth of the host, inclusion body (IB) formation, protein inactivity, and even not obtaining any protein at all are some of the problems that are experienced frequently. In this study, design of an expression system is discussed in detail in Chapter 5.

### 2.9. Conclusion

Bacterial cellulose, discovered over 130 years ago by Brown, has been gaining significant attention from scientists and studied extensively in literature. Genetic pathway has been investigated to identify to investigate the mechanism behind cellulose biosynthesis in bacteria. Different culturing conditions and tools have been used to increase the production. In this chapter, a research background is provided to support this study with relevant research evidence based on
literature and to obtain the aims of this study based on the demands in literature, which are discussed in the next chapter.
3. AIMS AND OBJECTIVES
3.1. Importance of Bacterial Cellulose for Industrial Applications

Bacterial cellulose (BC) is an environmentally-friendly polymer, which has been receiving increased attention in science and industry.

The unique properties of BC arise from its structure. It is an unbranched polymer with nanofibrils, made up of linear glucan chains, which has highly regular intra- and inter molecular hydrogen bonds. The hydrogen bond network makes cellulose a relatively stable polymer. This fine structure makes BC different from other microbial polysaccharides. BC has advantages over plant cellulose since it is crystalline in a pure form, which does not contain any lignin, pectin, or araban. The purification process of BC is extremely straightforward, which requires less energy [21, 80]. It also has specific chemical and mechanical properties, which are more advantageous for specific applications than plant derived cellulose. BC’s high purity, biocompatibility, tensile strength, crystallinity and hydrophilicity have all contributed to increasing interest in this biopolymer [29, 81, 82].

As BC is a highly pure material with unique properties, it has received attention as a green material with several industrial applications. BC nanofibres have a water-holding capacity up to 100 times their weight. In contrary to many synthetic polymers, BC is biocompatible and exhibits tissue integration. Thus, nanofibrils are able to guide cells. As a result of this, BC is very attractive as a scaffold for tissue engineering, as a replacement for skin tissue, cartilage, bone soft tissue blood vessel and cornea [83].

Moreover, its network in the structure allows it to incorporate drugs in order to support their controlled release [29, 84-90]. Furthermore, chronic wounds such as venous leg ulcers, or diabetic ulcers and burned tissues represent a significant clinical challenge, as they are problematic to heal. BC provides all requirements such as high porosity, non-toxicity, providing barrier against infection, for wound dressing material. Bioprocess®, XCell®, and Biofill® are products already available commercially for topical application in wound healing [24, 85]. Biofill
(Brazil) produced two products, Bioprocess and Gengiflex, as dressings for extensive wounds [57, 91]. Prima Cel™, produced by Xylos Corp (USA), has been applied in clinical tests to heal ulcers and wounds.

The significant dimensional stability of BC allowed the development of a sound transducing membrane, which is able to maintain high sonic velocity over a wide frequency ranges. This membrane produced of BC membranes satisfies the all rigid requirements for optimal sound transduction. Sony Corporation (Japan), in conjunction with Ajinomoto (Japan), developed audio speaker diaphragms using BC, however the product was not reasonable for the market because of high costs [92].

The production of high quality papers is also among the application area of BC, as it is highly flexible and durable [38]. Ajinomoto Corporation along with Mitsubishi Paper Mills in Japan has worked on the development of microbial cellulose for paper products [93]. The environmental damage caused by cutting down trees for the purpose of paper production and serious environmental pollutants released by burning of lignocellulosic products (e.g. paper and wood) results in a high risk for environment. When paper sheets produced by incorporating cellulose, they showed a flame-retardant behaviour [94, 95].

Also, its indigestibility in the human intestinal tract enables BC to be used in food industry as a multifunctional food ingredient to control the properties of food or beverage as a thickener, stabilizer and texture modifier. Nata, which is a BC product produced by Acetobacter xylinum, is a popular snack in the Philippines and in other countries. It has been used in diet drinks in Japan since 1992. Furthermore, Acetobacter was grown along with yeast in Manchurian tea extract and sugar for improved health purposes. It can also be used as a food packing material because of its edibility and biodegradability.
3.2. Research Challenges

*A. xylinum* is the model and the most commonly used microorganism for applied studies on cellulose, since it is source of bacterial cellulose which has ability to produce relatively high levels of BC from a wide range of carbon and nitrogen sources [11].

BC possesses an extreme potential for industrial applications in comparison with plant-derived cellulose and there are promising applications of bacterial cellulose in industry. However, these have not yet been fully exploited as it is still expensive compared with other popular commercial organic products. Published research data on BC productions and yields are summarized in Table 3.1. It is clear that productivity is still low for the wide use of BC in industry.

Currently plant-derived cellulose has a remarkable economic importance worldwide, which valued at $200 billion per year for cellulose market. In 2011, the price of wood pulp, the main source of plant-derived cellulose, reached its highest price in more than 30 years ($0.99 per kg). It was expected the global price of wood pulp to rise at an annualized rate of 1 percent in the two years to 2017, in consequent to a worldwide increase of industrial activities. The increasing demand of industrialization has raised environmental problems and influenced the ecological balance of our planet negatively. The extraction process of cellulose from wood accomplished via environmentally hazardous chemistries (bisulfites) as well as requires high energy [96]. To solve these problems, development of an ecologically friendly method of cellulose synthesis is required. One approach to limit the demand from plants is the production of cellulose using a bacterial system [96, 97]. The first headphones developed by Sony Corporation were extremely expensive, over $3,000 for one pair. Many recent developments have been published for cellulose-producing bacteria and the average cost has been achieved to decline to $77 per kg, which is 78 times higher than plant-derived cellulose [98].

BC productivity is mainly affected by culture conditions, such as the composition of the culture medium, pH, temperature, dissolved oxygen content
and the culturing strategy (static or agitated fermenters). The studies on improvement of BC productivity by optimisation of culture conditions have been summarized in previous chapter. Identification of genetic pathway of BC biosynthesis and application of genetic modification on *A. xylinum* have been also investigated, a mentioned in previous chapter. Although production of BC has been studied extensively, the improvement in BC production has not been satisfactory to reduce the high cost of BC.

**Table 3.1:** The BC productivities of various cellulose-producing bacteria. Adapted from Chawla et al. [99].

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Carbon source</th>
<th>Supplement</th>
<th>Culture time (h)</th>
<th>Cellulose Concentration (g L⁻¹)</th>
<th>Volumetric productivity (g L⁻¹h⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. xylinum</em> BRC 5</td>
<td>Glucose</td>
<td>Ethanol + oxygen</td>
<td>50</td>
<td>15.30</td>
<td>0.3060</td>
</tr>
<tr>
<td><em>G. hansenii</em></td>
<td>Glucose</td>
<td>Oxygen</td>
<td>48</td>
<td>1.72</td>
<td>0.036</td>
</tr>
<tr>
<td><em>G. hansenii</em></td>
<td>Glucose</td>
<td>Ethanol</td>
<td>72</td>
<td>2.50</td>
<td>0.035</td>
</tr>
<tr>
<td><em>Acetobacter sp.</em> V6</td>
<td>Glucose</td>
<td>Ethanol</td>
<td>192</td>
<td>4.16</td>
<td>0.022</td>
</tr>
<tr>
<td><em>Acetobacter sp.</em> A9</td>
<td>Glucose</td>
<td>Ethanol</td>
<td>192</td>
<td>15.20</td>
<td>0.079</td>
</tr>
<tr>
<td><em>A. xylinum</em> BPR2001</td>
<td>Molasses</td>
<td></td>
<td>72</td>
<td>7.82</td>
<td>0.109</td>
</tr>
<tr>
<td><em>A. xylinum</em> BPR2001</td>
<td>Fructose</td>
<td>Agar oxygen</td>
<td>72</td>
<td>14.10</td>
<td>0.196</td>
</tr>
<tr>
<td><em>A. xylinum</em> ATCC 53582</td>
<td>Glucose</td>
<td>Agar</td>
<td>56</td>
<td>12.00</td>
<td>0.215</td>
</tr>
<tr>
<td><em>A. xylinum</em> ATCC 53582</td>
<td>Glucose</td>
<td></td>
<td>168</td>
<td>2.73</td>
<td>0.016</td>
</tr>
<tr>
<td><em>A. xylinum</em> ssp. <em>sucrofermentans</em> BPR2001</td>
<td>Fructose</td>
<td>Oxygen</td>
<td>52</td>
<td>10.40</td>
<td>0.2</td>
</tr>
<tr>
<td><em>A. xylinum</em> ssp. <em>sucrofermentans</em> BPR2001</td>
<td>Fructose</td>
<td>Agar oxygen</td>
<td>44</td>
<td>8.70</td>
<td>0.198</td>
</tr>
<tr>
<td><em>A. xylinum</em> E25</td>
<td>Glucose</td>
<td></td>
<td>168</td>
<td>3.50</td>
<td>0.021</td>
</tr>
<tr>
<td><em>G. xylinus</em> K3</td>
<td>Mannitol</td>
<td>Green tea</td>
<td>168</td>
<td>3.34</td>
<td>0.019</td>
</tr>
<tr>
<td><em>G. xylinus</em> IFO 13773</td>
<td>Glucose</td>
<td>Lignosulphonate</td>
<td>168</td>
<td>10.10</td>
<td>0.060</td>
</tr>
<tr>
<td><em>A. xylinum</em> NUST4.1</td>
<td>Glucose</td>
<td>Sodium alginate</td>
<td>120</td>
<td>6.00</td>
<td>0.050</td>
</tr>
<tr>
<td><em>G. xylinus</em> IFO 13773</td>
<td>Molasses</td>
<td></td>
<td>168</td>
<td>5.76</td>
<td>0.034</td>
</tr>
<tr>
<td><em>Gluconacetobacter sp.</em> RKY5</td>
<td>Glycerol</td>
<td></td>
<td>144</td>
<td>5.63</td>
<td>0.039</td>
</tr>
</tbody>
</table>
3.3. Proposed Solution and Objectives of This Study

The challenge in large-scale production of a material by biological systems is the questionable nature of a living organism. Creating controlled and reproducible production techniques is essential for BC production to help way for greater acceptance of BC as the commercially available material.

Economic feasibility of BC is primarily dependent on BC productivity. The main challenge is the low productivity is correlated with the slow growth rate of *A. xylinum*, need to be overcome. Its doubling time is 4 hours in agitated and 8 hours in static cultures. Also, its susceptibility to the medium conditions has a significant influence on BC productivity. In order to overcome this, new valuable strains and technologies such as genetic modification should be explored.

On the other hand, *E. coli* is a well-characterised model organism which has been extensively studied. It can survive under various growth conditions in inexpensive culture media. Large scale production systems which are established on *E. coli* allow a rapid and economical production of desired products as its doubling time is approximately 20 minutes. BC production has been detected in some of the commensal and pathogenic stains of *E. coli* in a combination with other extracellular substances in microscopic level. For these reasons, *E. coli* displays a great potential to create a high-throughput production system.

As mentioned in the previous chapter, recombinant DNA technology is a very powerful technique to manipulate the host cells for the production of desired product. This could be achieved by cloning the target genes into an expression vector encoding the genes required for product of interest. In this study, a BC production system has been aimed to create by recombinant DNA technology based on *E. coli* in effort to generate valuable strains for improved BC production. For this purpose, the objectives of this study were:
• To design an expression system, compatible with different *E. coli* strains, in effort to investigate the production in various strains.

• To select and analyse the target fragments that encode cellulose biosynthesis, and optimise the fragments according to codon recognition of *E. coli*.

• To clone target fragments into expression vectors, and transform different *E. coli* strains with the recombinant vectors,

• To optimise bioprocess parameters by monitoring the growth profile and the plasmid stability at various temperatures and inducer concentrations,

• To verify the gene and protein expressions by SDS-PAGE and q-PCR analyses at various culture temperatures and inducer concentrations,

• To characterise the product by ATR-FTIR, scanning electron microscopy and light microscopy,

• To investigate the BC production in the presence of different concentrations of carbon sources,

• To design a fed batch system to culture genetically modified cells for improved BC production.
4. MATERIALS AND METHODS
4.1. MATERIALS

4.1.1. Bacterial Strains and Plasmids

In this study, *Acetobacter xylinum* ATCC 53582, *Escherichia coli* BL21 (DE3) HMS176 (DE3) and C41 (DE3) strains were used. *A. xylinum* ATCC 53582 was purchased from LGC Stardarts, *E. coli* strains were purchased from Novagen.

Expression system designed in our laboratory based on commercially available vectors, which are pETDuet-1 and pCDFDuet-1. Both vectors were purchased from Novagen. Target genes and promoters were synthesized, optimised and inserted into plasmids by Genewiz. pETDuet-1 was named as pBCS and pCDFDuet-1 was named as pCMP. *E. coli* strains and plasmid vectors that used in this study and their features are presented in Table 4.1.

**Table 4.1: E. coli strains and plasmid vectors that used in this study.**

<table>
<thead>
<tr>
<th><em>E. coli</em> Strain</th>
<th>Feature</th>
<th>Recombinant proteins synthesized</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> BL21 (DE3)</td>
<td><em>E. coli</em> BL21 (DE3) carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CcpAx</td>
</tr>
<tr>
<td><em>E. coli</em> HMS 176 (DE3)</td>
<td><em>E. coli</em> HMS176(DE3) carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CcpAx</td>
</tr>
<tr>
<td><em>E. coli</em> C41 (DE3)</td>
<td><em>E. coli</em> C41 (DE3) carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CcpAx</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Feature</th>
<th>Expressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBCS</td>
<td>pETDuet-1 with inserted <em>bcsABCD</em> operon encoding cellulose synthase</td>
<td><em>bcsA, bcsB, bcsC, bcsD</em></td>
</tr>
<tr>
<td>pCMP</td>
<td>pCDF Duet-1 with inserted <em>cmcAx</em> and <em>ccpAx</em> encoding endo-β-1,4-glucanase and cellulose complementing protein, respectively.</td>
<td><em>cmcax, CcpAx</em></td>
</tr>
</tbody>
</table>
4.1.2. Chemicals

All chemicals, solutions, and enzymes used in this study were purchased from APPLICHEM (Germany), MERCK (Germany), MOLEKULA (Germany) SIGMA (USA), PROMEGA (USA). Protein or DNA molecular weight markers were purchased from Fermentas (USA).

4.1.3. Growth Media for *Escherichia coli*

A nutrient rich media is required for high density cell growth for bacterial fermentation. Lysogeny broth (LB) media is widely used for recombinant protein expression. The ingredients used for the preparation of 1 litre of LB are shown in Table 4.2 [100].

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>10 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>5 g</td>
</tr>
<tr>
<td>NaCl</td>
<td>10 g</td>
</tr>
<tr>
<td>Add dH₂O up to 1 liter</td>
<td></td>
</tr>
</tbody>
</table>

The solids were suspended in 800 ml of distilled water. Further distilled water was added to reach a total volume of 1 liter. The media was autoclaved at 121 °C for 20 mins to sterilise.

The addition of agar to LB results in LB agar media, which is a media in the formation of a gel that bacteria can grow on (Table 4.3)[101]. Bacterial colonies unable to digest the agar but they can export nutrients from the LB agar media. The addition of an antibiotic to this media allows the selection of viable bacterial colonies only with the specific antibiotic resistance. This is usually conferred by a plasmid carrying the antibiotic resistance gene, which allows determination the plasmid stability.
Table 4.3: Composition of LB Agar Media.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>10 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>5 g</td>
</tr>
<tr>
<td>NaCl</td>
<td>10 g</td>
</tr>
<tr>
<td>Agar</td>
<td>15 g</td>
</tr>
<tr>
<td>Add dH₂O up to 1 liter</td>
<td></td>
</tr>
</tbody>
</table>

4.1.4. Growth Media for *Acetobacter xylinum* ATCC 53582

Hestrin-Schramm medium, HS, is the most preferable medium for the fermentation of bacterial cellulose producing bacteria [10]. HS medium consist of a small amount of carbon source, an enriched nitrogen source and a small amount of citric acid. The ingredients used for the preparation of 1 litre of HS and HS Agar Media are shown in Table 4.4 and Table 4.5, respectively.

Table 4.4: Composition of HS Media.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>5 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>5 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>20 g</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>2.7 g</td>
</tr>
<tr>
<td>Citric Acid</td>
<td>1.15 g</td>
</tr>
<tr>
<td>Add dH₂O up to 1 liter</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.5: Composition of HS Agar Media.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tryptone</td>
<td>5 g</td>
</tr>
<tr>
<td>Yeast extract</td>
<td>5 g</td>
</tr>
<tr>
<td>Glucose</td>
<td>20 g</td>
</tr>
<tr>
<td>Na₂HPO₄</td>
<td>2.7 g</td>
</tr>
<tr>
<td>Citric Acid</td>
<td>1.15 g</td>
</tr>
<tr>
<td>Agar</td>
<td>15 g</td>
</tr>
<tr>
<td>Add dH₂O up to 1 liter</td>
<td></td>
</tr>
</tbody>
</table>

4.1.5. Gels, Buffers, Solutions and Other Materials

Agarose gel (4% agarose dissolved in 40 mM TAE buffer) and TAE running buffer (40 mM Tris, 1 mM EDTA, pH 8.0) were used for Agarose Gel Analysis. NuPAGE Bis-Tris Gels and MOPS Running Buffer were purchased from Invitrogen for SDS-PAGE analysis. Compositions of fixing, washing and staining solutions for SDS-PAGE analysis is shown in Table 4.6, Table 4.7 and Table 4.8, respectively, according to the manufacturer’s instruction [102].

Table 4.6: SDS-PAGE Fixing Solution

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanol</td>
<td>50</td>
</tr>
<tr>
<td>Acetic acid</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 4.7: SDS-PAGE Washing Solution

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanol</td>
<td>34</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>17</td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.8: SDS-PAGE Staining Solution

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Amount (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanol</td>
<td>34</td>
</tr>
<tr>
<td>Ammonium sulfate</td>
<td>17</td>
</tr>
<tr>
<td>Phosphoric acid</td>
<td>2</td>
</tr>
<tr>
<td>Commassie Brilliant Blue - G250</td>
<td>0.066</td>
</tr>
</tbody>
</table>

For protein extraction, Bugbuster extraction solution was purchased from Navogen. RNAprotect Bacteria Reagent was used to stabilize RNA immediately prior to RNA isolation procedure. Custom primers were purchased from Invitrogen.

4.1.6. Kits

The kits purchased for experimental analysis in this study are shown in Table 4.9.

4.1.7. Other Stocks and Solutions

- **1 M IPTG Stock**: 2.38 g IPTG was dissolved in 10 ml distilled water and the solution was sterilized with 0.22μm filter.
- **Carbenicillin Stock (50 mg/ml)**: 0.5 g carbenicillin was dissolved in 10 ml distilled water and the solution was sterilized with 0.22μm filter.
- **Streptomycin Stock (50 mg/ml)**: 0.5 g streptomycin was dissolved in 10 ml distilled water and the solution was sterilized with 0.22μm filter.

**Table 4.9**: Kits used in this study.

<table>
<thead>
<tr>
<th>Kit</th>
<th>Application</th>
<th>Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takara Long Range PCR Kit</td>
<td>PCR</td>
<td>RR013A</td>
</tr>
<tr>
<td>Qiagen RNeasy Mini Kit</td>
<td>RNA Extraction</td>
<td>74104</td>
</tr>
<tr>
<td>QuantiTect Reverse Transcription Kit</td>
<td>Reverse Transcription</td>
<td>205310</td>
</tr>
<tr>
<td>QuantiTect SYBR® Green RT-PCR Kit</td>
<td>Quantitative PCR</td>
<td>204243</td>
</tr>
<tr>
<td>Sigma Glucose Assay Kit</td>
<td>Determination of free Glucose Concentration</td>
<td>G3293</td>
</tr>
<tr>
<td>Sigma Free Glycerol Determination Kit</td>
<td>Determination of free Glycerol Concentration</td>
<td>FG0100</td>
</tr>
<tr>
<td>Acetate Colorimetric Assay Kit</td>
<td>Determination of free Acetate Concentration</td>
<td>MAK086</td>
</tr>
</tbody>
</table>
4.2. METHODS

4.2.1. Sterilization

In this study, sterilized equipments were used and experimental work that required sterile environment was carried out. Equipment including glassware, tips, tubes, solutions and culture media and supplements were sterilized in autoclave at 1.02 atm and 121°C for 15 minutes. Antibiotics, glycerol, glucose and IPTG and other solutions used for fermentation were filter sterilized using sterilized membranes with a pore size of 30 μm.

4.2.2. Preparation of Bacterial Stocks

A single colony from a master plate with antibiotic supplement (if necessary) was inoculated with a sterile tip into fresh 5 ml LB medium supplemented with appropriate antibiotic and grown overnight at 37 °C and 180 rpm. 500 μL grown cells were added to sterile tubes containing 30 % sterile glycerol and stocks were stored at -80°C.

4.2.3. Preparation of Preculture

4.2.3.1. E. coli Strains

5 ml of sterile liquid medium supplemented with the appropriate antibiotic (if necessary), was inoculated with 50 μL frozen stock culture.

4.2.3.2. A. xylinum Strains

Pre-cultures were prepared in 50 mL flasks containing 30 mL of sterile HS medium at 30°C in a water bath operating at 25 rpm. Following 48 hours of culture, inocula were prepared by 1:10 dilution with fresh sterile HS medium.
4.2.4. Growth Conditions

4.2.4.1. *E. coli* Strains

The cell growth experiments were conducted either in flasks or in stirred-tank bioreactor. Orbital shaker was operated 180 rpm for the flask experiments unless otherwise noted. Culture volume was kept at one fifth of the flask volume to provide efficient aeration. The various temperature values (22°C, 30°C, 37°C) were used to monitor the growth. The bioreactor was operated at 250 rpm with 20% O2 supplement at 22°C or 30°C. pH was set at 7.0±0.5.

LB medium supplemented with glycerol or glucose was used as growth media in experiments. The growth media were supplemented with 50 μg/ml carbenicillin and 50 μg/ml streptomycin for *E. coli* cells harbouring pETDuet-1 and pCDFDuet-1. The media was supplemented by various concentrations of IPTG (isopropyl β-D-1-thiogalactopyranoside) for induction of the expression of the target *bcs* genes. IPTG was added to the growing culture when OD$_{600}$ of cell culture reached 1 unless otherwise noted.

4.2.4.2. *A. xylinum* Strains

Cells were grown in orbital shakers at 30°C and 25 rpm. Culture volume was kept at three fifth of the flask volume for aeration.

4.2.5. Monitoring Cell Growth

Optical densities of cells were measured at 600 nm using spectrophotometer to monitor cell growth. The growth medium for *A. xylinum* was supplemented by 0.1% cellulase in order to have a uniform culture. Samples were diluted to keep the spectroscopic readings between reliable limits (0.2-0.8).

Optical density analysis is based on Lambert-Beer law which represents the direct and linear relationship between the amount of light adsorbed and the concentration of constituents in the sample.
Although measuring optical density is a straightforward method, it cannot provide information about the amount of viable cells in the culture. This can be accomplished by serial diluting the bacteria and plating the diluted bacteria on media supplemented by agar. This method is more time consuming, however, it provides statistically accurate and repeatable results.

4.2.6. Viable Cell Count and Plasmid Stability Test

Bacterial samples were plated to determine the number of total viable cells and the number of viable cells harbouring plasmids following induction with IPTG to estimate plasmid stability. For a ten-fold dilution, 100 μl of the bacterial sample taken from the growing cell suspension at the desired time interval was transferred into a glass tube containing 0.9 ml of fresh liquid medium and mixed by vortex to provide homogenization. For another ten-fold dilution, 100 μl from this mixture was taken and deposited in a new glass tube containing 0.9 ml fresh liquid medium. This process continued until the desired dilution was reached, as presented in Figure 4.1. Finally, 1 ml of the dilution was deposited in a sterile Petri dish and agar containing appropriate antibiotics (carbenicillin and streptomycin) to determine the amount of viable cells harbouring plasmids. In order to determine the amount of total viable cells, 1 ml of the dilution was deposited in a sterile Petri dish and agar without any antibiotic supplement. Petri dish was mixed gently by moving it through an eight-shaped path on the ground without lifting. The plates were left to cool to harden and then incubated overnight at 37°C in an incubator. The colonies were counted to determine the colony forming units on each Petri dish. Average values from triplicates plates were taken.
4.2.7. Plasmid Transformation into *E. coli* Cells

Transformation is the process of introducing the recombinant vector from a reaction mixture or vector solution into *E. coli* cells. This can be achieved by chemical transformation or electroporation. In this study, chemical transformation was used for the transformation of *E. coli* cells (Figure 4.2).

Competent cell tubes were removed from the freezer and allowed to thaw on ice for 2–5 min. The tubes were gently flicked 1–2 times to evenly resuspend the cells. 20 μl aliquots of cells were pipetted into pre-chilled tubes. 1 μl of (1-2 ng) plasmid DNA added directly to the cells. The tubes were stirred gently to mix and retured to the ice, making sure that the tube is surrounded by ice except for the cap. Tubes were incubated on ice for 30 minutes to allow the plasmid DNA to enter the cells. Following a heat shock at 42 °C for exactly 30 seconds, the cells were further incubated on ice for ten more minutes. 80 μl of room temperature SOC added to each tube without any antibiotics. The tubes were kept on ice until all have received SOC and then the cells were incubated at 37 °C for one hour at 250 rpm. The transformed cells were centrifuged at 3000 rpm for five minutes. Selection for transformation was accomplished by plating the cells on medium.
containing antibiotics (25 mg/ml carbenicillin and 25 mg/ml streptomycin) for the plasmid encoded drug resistance. The plates were incubated overnight at 37 °C.

Determination of Dry Cell Weight

Reliable estimates of bacterial biomass are essential for quantitative analysis. For this purpose, cells were grown overnight at 180 rpm and 37 °C in 100 ml of medium in a 500 ml of flask. Cells were centrifuged at 9000 rpm and washed with distilled water for two to three times. 50 ml of the cells, resuspended in 100 ml of distilled water, were dried at 80 °C and the dry cell weight was determined. The
remaining 50 ml, with the optical density 0.4-0.6 was used to prepare diluted samples containing 0.125, 0.250, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5 and 5 ml original cell culture. All volumes were completed to 5 ml by distilled water. OD$_{600}$ nm values were read in a spectrophotometer. Cell dry weight (mg/ml) for each sample was determined using data from the dried cell sample. Calibration chart was prepared by plotting optical density vs. cell dry weight, as presented in Appendix 2. The slope of this graph was used to determine the dry cell weight of a known OD at 600 nm.

4.2.9. Extraction of Proteins

4.2.9.1. Extraction of periplasmic proteins

Periplasmic proteins were extracted from cells by osmotic shock procedure which was described by Nossal et al. 100 ml of sterile LB medium was inoculated with a preculture at a final OD of 0.7 and placed in an orbital shaker at 37°C, 180 rpm. In cases where induction with IPTG was needed, E. coli cells harboring pBCS and pCMP were grown until the optical density of the cell culture reached 1.0 at 600 nm and IPTG was added. Cells were allowed to grow 5 hours or 24 hr after induction. Grown cells were centrifuged for 10 min at 7,000 rpm, +4°C and the supernatant was discarded. After harvesting, cells were resuspended in 20 % sucrose solution (w/w) in 30 mM Tris-HCl (pH 8.0) with 1 mM ethylenediaminetetraacetic acid (EDTA). Cells were gently incubated for 20 min at room temperature. Following centrifugation at 9,000 rpm for 20 min, the pellet was rapidly resuspended in ice-cold 5 mM MgCl$_2$ and gently incubated for 20 min. Cells were removed from the periplasmic protein extract by centrifugation at 9,000 rpm, 4°C for 20 min. Supernatant contained periplasmic proteins.

4.2.9.2. Extraction of cytoplasmic proteins

The pellet from periplasmic fraction was resuspended in room temperature BugBuster by pipetting or gentle vortexing using 5 ml reagent per gram of wet cell paste. Typically, 2 ml of BugBuster is used per 50 ml of culture. 1 KU rLysozyme was added solution per 1 ml BugBuster Reagent (5 KU/g cell paste). Following, the incubation of cell suspension on a shaking platform or rotating mixer at a slow
setting for 30 min at room temperature, the suspension was centrifuged at 16,000 × g for 20 min at 4°C to remove insoluble cell debris. The pellet was used to isolate the insoluble cytoplasmic fraction. The supernatant was transferred to a fresh tube for analysis and mixed a sample of the supernatant with an equal volume of 4X SDS Sample Buffer (Navogen). The sample was immediately heated for 3 min at 85°C to denature protein and then store at 20°C until SDS-PAGE analysis.

4.2.9.3. Extraction of insoluble fractions (inclusion bodies)

The insoluble cytoplasmic fraction may consist of cell debris, membrane proteins and aggregated target protein known as inclusion bodies. Inclusion bodies can be further purified by repeated centrifugation and washing steps. However, the product will be contaminated at some level with other proteins and nucleic acids. In some cases, purified inclusion bodies are suitable for direct use as antigens for the preparation of antibodies against the target protein. Some target proteins associated with the insoluble fraction may not be in inclusion bodies. Membrane associated target proteins can pellet with the insoluble fraction and may be released into the soluble fraction by including a detergent during lysis.

Using the insoluble pellet from the extraction of cytoplasmic proteins step, the pellet was resuspended in the same volume of 1X BugBuster Reagent that was used to resuspend the original cell pellet. In order to obtain an even suspension, the sample was pipetted up and down and vortexed. Complete resuspension of the pellet is critical to obtaining a high purity preparation in order to solubilize and remove contaminating proteins.

rLysozyme™ solution was added to a final concentration of 1 KU/ml. Following vortexing gently to mix, the sample was incubated at room temperature for 5 min. Then 6 volumes of 1:10 diluted BugBuster Reagent (in deionized water) was added to the suspension and the sample was vortexed again for 1 min. the suspension was centrifuged at 5,000 × g for 15 min at 4°C to collect the inclusion bodies. After resuspension of the inclusion bodies in ½ the original culture volume of 1:10 diluted BugBuster, the sample was again mixed by vortex, and centrifuged
as in previous step. The supernatant was removed and this step was repeated again. Resuspend once more but centrifuge at 16,000 × g for 15 min at 4°C and remove the supernatant.

For SDS-PAGE analysis, the final pellet was resuspended in 1.5 ml 1% SDS Sample Buffer with heating and vigorous mixing. 100 μl of sample was removed and combined with 100 μl of 4X SDS Sample Buffer (Navogen). Following this step, the sample was immediately heated for 3 min at 85°C to denature proteins and then store at –20°C until SDS-PAGE analysis.

4.2.9.4. Extraction of total proteins

PopCulture Reagent efficiently extracts protein from E. coli directly in the culture medium without cell harvest. Using this method, cell culture, protein extraction and purification can all performed in the original culture tube, flask or multiwell plate.

An induced culture of E. coli was treated with 0.1 culture volume of PopCulture Reagent for 10–15 min at room temperature in the presence of 40 U (1 μl of a 1:750 dilution) rLysozyme per 1 ml of original culture volume with treatment of Benzonase Nuclease. The suspension was pipetted up and down to mix and incubate for 1 hr at room temperature. This prepared extract can be analysed directly by SDS-PAGE of total cell protein by combining an aliquot of the prepared extract with an equal volume 4X SDS Sample Buffer (Cat. No. 70607-3) for detection with Coomassie staining.

4.2.9.5. Determination of Protein Concentration

Protein concentrations of the extracted proteins were determined using the method described by Bradford (1976). Samples containing protein were mixed with Bradford solution and their absorbance at 595 nm was recorded. Bovine Serum Albumin (BSA) was used as the standard for the preparation of the calibration curve. Samples with different BSA concentrations were and mixed with Bradford solution. The absorbance of these samples at 595 nm was measured to construct
the plot that relates protein concentration to the absorbance at 595 nm. The calibration curve is given in Appendix 1.

4.2.9.6. SDS-PAGE Analysis of Proteins

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a technique that involves the separation of proteins based on their sizes in an electric field. It is one of the most common analytical techniques used to separate and characterize proteins. Sodium dodecyl sulfate (SDS) is an amphipathic detergent, which binds non-covalently to proteins. In the presence SDS, proteins are disassociated from each other, and the intrinsic charge of a protein is masked. During SDS-PAGE analysis, all proteins migrate toward the anode.

The gel is constitutes of acrylamide and bisacrylamide. The 'pore size' is determined by the ratio of acrylamide to bisacrylamide. A high ratio of bisacrylamide to acrylamide and a high acrylamide concentration cause low electrophoretic mobility. Proteins with high molecular weight immigrates the gel slower than the proteins with low molecular weight. Based on this phenomenon, it is possible to determine the size of a denaturised protein by comparing with a molecular weight marker.

For this purpose, protein samples were mixed with 2X Sample Buffer containing SDS at a 1:1 ratio and boiled for 5 minutes to denature the proteins. The marker and the samples were loaded into the wells of 12 % SDS polyacrylamide gels. Proteins were allowed to run at 120 V for approximately 90 minutes. Then gels were gently removed from glass plates and placed in fixing solution for 30 mins to fix the proteins into the gel. After fixing, gels were washed with washing solution for 30 mins. Finally, gels were stained with Commasie G-250 at least overnight and destaining solution was used to remove excess unbound stain. The gels were dried for 40 min at 80 °C. The protein bands of the samples were compared against the marker to verify the presence of target proteins.
4.2.10. DNA Isolation

Bacteria are lysed and proteins are removed by digestion with proteinase K. Cell wall debris, polysaccharides, and residual proteins are removed by selective precipitation with cetyltrimethylammonium bromide (CTAB), and purification is completed by a phenol/chloroform extraction. Finally, DNA is recovered from the resulting supernatant by isopropanol precipitation.

Cells were grown until the end of the exponential growth phase. 1.5 mL of the culture was centrifuged for 2 min at high speed and the supernatant was discarded. The pellet was resuspended in 567 μl TE buffer. Following the addition of 10 μl lysozyme (10 mg/mL), the samples were mixed thoroughly and incubated for 1 hr min at 37°C. 30 μl of SDS and 3 μl of proteinase K were added and the suspension was mixed thoroughly and incubated for 1 h at 37°C. 100 μl of 5 M NaCl were added and mix thoroughly. After addition of 80 μl of CTAB/NaCl solution, the sample was mixed thoroughly, and incubated for 10 min at 65°C. 750 μl of chloroform/isoamyl alcohol was added and the sample was spined 10 min at high speed at room temperature.

After centrifugation, a white interface, the CTAB-protein/polysaccharide complexes to be eliminated, was visible. The viscous aqueous supernatant was collected carefully in a fresh microcentrifuge tube and left behind the interface. An approximately equal volume (400-600 μl) of phenol/chloroform/isoamyl alcohol was added and the sample was mixed thoroughly (minimum 10 s/tube), and spin as in step 5. This phenol/chloroform extraction step completes the elimination of proteins and CTAB precipitates. The supernatant was collected carefully in a fresh microcentrifuge tube. 0.6 vol isopropanol was added to precipitate the nucleic acids. The sample was mixed from top to bottom until a stringy whitish DNA precipitate is clearly visible. After the spinning 10 min at high speed in a microfuge at room temperature, the supernatant was eliminated. The position of the DNA pellet was verified before removing the supernatant. After that, the DNA pellet of was washed with 500 μl 70% ethanol and spinned as in the previous steps. The supernatant was carefully removed with a pipette and the pellet was left to dry at
room temperature. The DNA was pelleted in 30 μl DNase free water. The concentration of DNA and the purity of sample were checked by Nanodrop. DNA sample was stored at -20°C for further analysis.

4.2.11. RNA Isolation

For an accurate analysis of gene expression in bacteria, it is important to analyse RNA that truly represents in vivo gene expression. If conventional methods are used to harvest bacterial cells, 2 major problems may affect the gene expression profile. Firstly, RNA may be enzymatically degraded, and this results in a loss of many transcripts. Since bacterial mRNAs usually only have a very short half-life (only a few minutes), this is particularly significant for the extraction process. The other problem is that genes may be induced during handling of bacterial cells, leading to higher expression of specific genes. In this study RNAProtect Bacteria Reagent was used to stabilize RNA before bacterial cells are lysed. This reagent prevents both degradation of RNA transcripts and induction of genes to ensure reliable gene expression analysis. After cell lysis, the RNeasy Mini Kit was used to purify total RNA. The resulting high-quality RNA reflects the accurate in vivo gene expression profile of bacteria and is suitable for use in quantitative PCR.

In this protocol, bacterial lysis was achieved by treatment with 1 mg/ml lysozyme for 5 min, which is optimal for E. coli and RNA collected by the binding capacity of the RNeasy spin column. First, 1 volume of bacterial culture was mixed with 2 volumes of RNAProtect Bacteria Reagent. Following the mixing by vortex immediately for 5 s, the culture was incubated in room temperature for 5 min. After centrifugation for 5 min at 5000 x g, the supernatant was discharged. In this step, the cell pellet can be store -20°C for up to 2 weeks or -70°C for up to 4 weeks. 200 μl of TE buffer (10 mM Tris·HCl, 1 mM EDTA, pH 8.0) containing 1 mg/ml lysozyme was used for cell lysis. The sample was mixed by vortex for 10 sec and incubated at room temperature for 30 min. Following the addition of 700 μl of RLT buffer containing β-mercaptoethanol, sample was mixed by vortex again. After
that, 500 μl of ethanol added and the sample was pipetted to have a uniform mixture, as a precipitation may form at this step.

Following the cell lysis step, mini spin column was used for the purification of RNA. Up to 700 μl lysate (half of the sample) was transferred, including any precipitate that may have formed, to a spin column placed in a 2 ml collection tube and centrifuged for 15 s at 8000 x g (10,000 rpm). After discarding the flowthrough, this step repeated again for the rest of the lysate. In order to on-column DNase digestion during RNA purification, DNase I stock solution was prepared by dissolving the solid DNase I (1500 Kunitz units) in 550 μl of the RNase free water. 350 μl of Buffer RW1 was added to the spin column and centrifuged for 15 s at 8000 x g. After discard the flow-through, 10 μl DNase I stock solution added to 70 μl Buffer RDD to prepare DNase I incubation mix. DNase I incubation mix was added directly to the RNeasy spin column membrane. After incubation at room temperature (20–30°C) for 15 min, 350 μl Buffer RW1 added to the RNeasy spin column. Following 5 min incubation, the column was centrifuged for 15 s at ≥8000 x g and the supernatant was discharged.

After DNA digestion step, the spin column placed into a new 2 ml collection tube and 500 μl Buffer RPE buffer was added. Following centrifugation for 15 s at ≥8000 x g (≥10,000 rpm) to wash the spin column membrane, the flow-through was discharged and 500 μl Buffer RPE was added to the spin column. The column was centrifuge for 15 s at ≥8000 x g (≥10,000 rpm) to wash the spin column membrane. As a final step the spin column was placed in a new 1.5 ml collection tube and 30 μl RNase-free water was added directly to the spin column membrane. The column was centrifuged for 1 min at ≥8000 x g (≥10,000 rpm) to elute the RNA. The concentration of DNA and the purity of sample were checked by NAnodrop. DNA sample was stored at -20°C for further analysis.
**4.2.12. Polymerase Chain Reaction**

Polymerase chain reaction (PCR) is an efficient and cost-effective enzymatic assay that enables in vitro amplification of a specific DNA fragment from a complex pool of DNA.

PCR is based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the template strand in the existence of primer to initiate the amplification.

Each PCR assay requires the presence of template DNA, primers, nucleotides, DNA polymerase and a buffer. DNA polymerase is a heat resistant enzyme that can generate new DNA strands. DNA template that contains the target sequence that needs to be amplified.

Primers are short pieces of single-stranded DNA that are complementary to target sequence. They take part in the specification of target sequence and the initiation of the replication of DNA. Nucleotides (dNTPs) are single units of bases (adenine, thymine, cytosine, and guanine) that are essential to build new strands. Buffer (usually MgCl₂) is a salt-solution that keeps the DNA and other components stabilized.

After the components mixed in a test tube, they placed in a thermocycler to reach required temperature values for each cycle. First, the reaction mixture is heated above the melting point of the two complementary DNA strands of the target DNA, in order to separate the strands. This step is called denaturation. The temperature is then decreased to a specific degree for primers to bind to the target DNA segments, a step called annealing. After that, the temperature is raised again, in order to allow DNA polymerase to extend the primers by adding nucleotides to the developing DNA strand. This step is known as extension (Figure 4.3). These three steps are repeated for 30-40 times to generate sufficient copy of target DNA sequences.
Takara LA PCR kit was used since it allows amplifying long PCR products accurately. DNA sequences longer than 5 kbp is difficult to amplify efficiently and this limits the wide range of application of PCR. Another limitation to the PCR method is the incorporation of mismatched nucleotides. Recently, conditions such as DNA polymerase, reaction buffers, and thermal cycling profiles were improved for accurate amplification of longer DNA up to 40 kbp. Takara LA PCR technology has been selected for PCR analysis in this study.

The reaction mixture was prepared in reaction tube by combining the following reagents, Table 4.10, to the total volume of 50 μl. G-Storm Thermocyler System was used to run the PCR samples. Reaction steps and durations are shown in Table 4.11 The samples were stored at -20°C after the cycle was completed. Agarose gel electrophoresis was used to analyse the PCR samples.
Table 4.10: Reaction mixture for PCR.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μl)</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA Taq Polymerase</td>
<td>0.5</td>
<td>2.5 units/50 μl</td>
</tr>
<tr>
<td>10X LA PCR Buffer II (Mg²⁺)</td>
<td>5</td>
<td>1X</td>
</tr>
<tr>
<td>DNA template</td>
<td>1</td>
<td>100ng/50 μl</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>0.5</td>
<td>0.5 μM</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>0.5</td>
<td>0.5 μM</td>
</tr>
<tr>
<td>dNTP</td>
<td>8</td>
<td>400 μM</td>
</tr>
<tr>
<td>H₂O</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Total Volume</td>
<td>50 μl</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.11: Reaction steps for PCR.

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial denaturation</td>
<td>94°C</td>
<td>1 min</td>
</tr>
<tr>
<td>Denaturation</td>
<td>94°C</td>
<td>30 sec</td>
</tr>
<tr>
<td>Cycles (30-40)</td>
<td>Annealing</td>
<td>65°C</td>
</tr>
<tr>
<td></td>
<td>Extension</td>
<td>72°C</td>
</tr>
<tr>
<td></td>
<td>Final extension</td>
<td>72°C</td>
</tr>
</tbody>
</table>
4.2.13. Primer Design

Primers used in this study were designed by NCBI/Primer-BLAST tool. The properties of the primers is summarised in Table 4.12. For Real Time q-PCR, as short DNA sequences (usually 100 to 600 bases) can be amplified; primers design to amplify an approximately 100 bases from beginning of the target gene including T7lac promoter and an approximately 100 bases from end of the same target gene including stop codon.

<table>
<thead>
<tr>
<th>Oligo sequence (5' to 3')</th>
<th>Oligo name</th>
<th>Length</th>
<th>Tm</th>
<th>GC%</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGCGACTCTCCGTATTAGGAAATTA</td>
<td>cmc1 (F)</td>
<td>24</td>
<td>60.4</td>
<td>42</td>
</tr>
<tr>
<td>TGCGATCCATCACACTCATCG</td>
<td>cmc1 (R)</td>
<td>21</td>
<td>63.0</td>
<td>57</td>
</tr>
<tr>
<td>TGTACACGGCCGATAATCG</td>
<td>ccp1 (F)</td>
<td>20</td>
<td>60.9</td>
<td>55</td>
</tr>
<tr>
<td>CACTACCCGCTGGCGCTCAT</td>
<td>ccp1 (R)</td>
<td>19</td>
<td>62.7</td>
<td>63</td>
</tr>
<tr>
<td>TCCCGCCAAATTATAACGACTCAT</td>
<td>A1 (F)</td>
<td>25</td>
<td>62.4</td>
<td>44</td>
</tr>
<tr>
<td>ATCGGATCTTGCTGTTGGTA</td>
<td>A1 (R)</td>
<td>21</td>
<td>63.3</td>
<td>57</td>
</tr>
<tr>
<td>TGCCAGTGCCCCGGTGTAAT</td>
<td>B1(F)</td>
<td>19</td>
<td>58</td>
<td>59</td>
</tr>
<tr>
<td>GTGCGCTTGGCTGCTGTCG</td>
<td>B1(R)</td>
<td>19</td>
<td>63</td>
<td>63</td>
</tr>
<tr>
<td>ACACCGCCGATAATCGAAAT</td>
<td>C1 (F)</td>
<td>21</td>
<td>61</td>
<td>48</td>
</tr>
<tr>
<td>GCGTAGCGCTTGTTGTGTCAT</td>
<td>C1 (R)</td>
<td>20</td>
<td>62</td>
<td>55</td>
</tr>
<tr>
<td>TGATTAGCGCCCCACTACCTGA</td>
<td>D1 (F)</td>
<td>21</td>
<td>61</td>
<td>52</td>
</tr>
<tr>
<td>GGAACAGGGTGAAATCGCGGC</td>
<td>D1 (R)</td>
<td>20</td>
<td>63</td>
<td>65</td>
</tr>
<tr>
<td>TGCCAGAATGCACCGGTTTGA</td>
<td>cmc1 (F)</td>
<td>22</td>
<td>63</td>
<td>50</td>
</tr>
<tr>
<td>TTACTTGTAGTTTTCTGGCCAGC</td>
<td>cmc1 (R)</td>
<td>24</td>
<td>63</td>
<td>50</td>
</tr>
<tr>
<td>GTTGCAAGGCCCTAGGGTAATC</td>
<td>ccp2 (F)</td>
<td>22</td>
<td>64</td>
<td>59</td>
</tr>
<tr>
<td>CGCAGTAATGCTACGCAGAC</td>
<td>ccp2 (R)</td>
<td>22</td>
<td>64</td>
<td>59</td>
</tr>
<tr>
<td>AGTGTTAGTTCCGTTGGTTCG</td>
<td>A2 (F)</td>
<td>22</td>
<td>63</td>
<td>55</td>
</tr>
<tr>
<td>TTTAAGCGCGAAGCCGACCTAGC</td>
<td>A2 (R)</td>
<td>22</td>
<td>63</td>
<td>55</td>
</tr>
<tr>
<td>GTTGCGCTCGGGTTTGTA</td>
<td>B2 (F)</td>
<td>20</td>
<td>63</td>
<td>60</td>
</tr>
<tr>
<td>TTAGCTTTACGGCGCCTGCTCC</td>
<td>B2 (R)</td>
<td>22</td>
<td>64</td>
<td>59</td>
</tr>
<tr>
<td>GTGCGGCGGTGTATCGGT</td>
<td>C2 (F)</td>
<td>20</td>
<td>61</td>
<td>55</td>
</tr>
<tr>
<td>TTATGATCATGATCGTAGTG</td>
<td>C2 (R)</td>
<td>25</td>
<td>58</td>
<td>40</td>
</tr>
<tr>
<td>GGCTGTATGCTGTTGGGT</td>
<td>D2 (F)</td>
<td>22</td>
<td>62</td>
<td>55</td>
</tr>
<tr>
<td>TGATGATCGCTGACCAGCCAC</td>
<td>D2 (R)</td>
<td>22</td>
<td>65</td>
<td>59</td>
</tr>
<tr>
<td>GCCGGGTTAAGGCAATCGGT</td>
<td>serC (F)</td>
<td>20</td>
<td>65</td>
<td>60</td>
</tr>
<tr>
<td>GCCGAAGTCGCGTCTCGT</td>
<td>serC (R)</td>
<td>20</td>
<td>65</td>
<td>65</td>
</tr>
</tbody>
</table>
4.2.14. Agarose Gel Electrophoresis

Agarose gel electrophoresis is a method to separate and visualize DNA fragments based on their size. The method is based on the fact that DNA is negatively charged at neutral pH due to its phosphate backbone. When an electrical potential is placed on negatively charge molecule, they migrate toward anode (Figure 4.4). The migration speed is depends on the size of molecules. Molecules with small weight migrate faster than larger ones, since they can fit though the holes in agarose gel easier [101].

Nucleic acid in agarose gel fragments are visualized on a U.V. trans-illuminator following an incubation of the gel in a florescent dye. The most commonly used dyes are ethidium bromide or SYBR Green.

In this study following method was used for agarose gel electrophoresis. 0.5 g of agarose was placed in a 250 ml glass flask. 50 ml TAE buffer was added to the flask. Agarose solution was heated in order to dissolve the agarose completely in the buffer for 1 minute in microwave. The solution was allowed to cool to a safe temperature to handle. After that, agarose solution was poured into gel trays. The
gel comp was placed into the poured agarose and any air bubbles were removed from the solution with a pipette tip. The gel was allowed to set at room temperature for approximately 20 minutes. Once the gel is set, the comb was removed slowly by lifting. After that, the gel tray was placed in the gel electrophoresis tank slowly. Once the gel is in the tank, the PCR samples can be loaded into the gel lanes. The samples were mixed with the required volume of 2x gel loading buffer and loaded on gel. Also the molecular weight marker was loaded to enable size estimation of bands on gel. The power supply was set to required 90 ampers to run the samples the length of the gel. The gel was visualized by a U.V. trans-illuminator.

4.2.15. Reverse Transcription Procedure

RNA is a highly unstable single-strand compared to DNA. Following the RNA purification step, RNA samples converted into cDNA immediately by reverse transcription procedure.

Reverse transcription is a process of generating single-stranded DNA (complementary DNA, or cDNA) in vitro using single-stranded RNA as a template in the presence of reverse transcriptases (RTs). The cDNA can be used as a template for amplification by quantitative RT-PCR. Principle of reverse transcription procedure is the same as in PCR, as presented in Figure 4.5. It is a mid-step between RNA extraction and q-PCR to generate cDNA from RNA samples in order to quantify the expression of desired DNA fragments.

In this study, Quantitech Reverse Transcription kit was used for reverse transcription. The composition of the reaction mixture can be seen in the following Table 4.13.
The mixture was incubated for 15 min at 42°C in thermocycler. After that, the mixture was incubated for 3 more min at 95°C in order to inactivate Quantiscript Reverse Transcriptase. For long-term storage, the reverse-transcription reaction was stored at −20°C. The steps of the procedure are presented in Figure 4.8.
4.2.16. Real Time q-PCR

Real-time PCR is a powerful tool that allows quantification of exponential amplification of short DNA sequences (usually 100 to 600 bases) in a complex mixture. The amount of final amplified product can be quantify by real-time PCR precisely and sensitively, at the end-point even if the starting amount of material is at a very low concentration [103-105]. It was first used in 1992 by Higuchi. Real-time PCR monitors the fluorescence emitted during the reaction as an indicator of amplification production during each PCR cycle (in real time) as opposed to the endpoint detection. The amount of genetic material present in the beginning correlates with how quickly the amplified target reaches a threshold detection level. The quantitative endpoint for real-time PCR is the threshold cycle (CT). The CT is defined as the PCR cycle at which the fluorescent signal of the reporter dye crosses an arbitrarily placed threshold. By presenting data as the CT, one ensures that the PCR is in the exponential phase of amplification. The numerical value of the CT is inversely related to the amount of amplification in the reaction. (i.e., the lower the CT, the greater the amount of amplification)[106-108].

Quantification of gene expression levels can generate meaningful information about gene function. The type of cells or tissue where a gene is expressed can be detected by an accurate measurement of gene expression. Biological state (e.g., disease, development, differentiation) can be defined and alteration in gene expression levels in response to specific biological stimuli (e.g., growth factor or pharmacological agent) can be detected. Other methods to study gene expression such as northern blot analysis and RNase protection assays are time consuming and sometimes prove difficult to reliably generate accurate measurements of gene expression levels. The use of real-time RT-PCR provides several advantages over these methods, including the relatively small amount of sample required for analysis, the ability to reproduce rapid and accurate data, and the capacity for analysing more than one gene at a time.
Although there are many different variations of real-time RT-PCR; SYBR Green is the most common approach that has been using. In this study, QuantiTect SYBR Green PCR Kit was used for Real Time q-PCR analysis. The composition of reaction mixture was prepared as shown in Table 4.14 and the run method was set up as shown in Table 4.15. Applied Biosystems Real Time PCR machine was used to run the PCR reactions in 96 well plates. The machine generates amplification plots during a run, which is presented in Figure 4.9.

**Table 4.14: Reaction steps for PCR.**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume/reaction</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x QuantiTect SYBR Green</td>
<td>10 μl</td>
<td>1x</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>0.5 μl</td>
<td>0.3 μM</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>0.5 μl</td>
<td>0.3 μM</td>
</tr>
<tr>
<td>Template cDNA</td>
<td>1 μl</td>
<td>100 ng/reaction</td>
</tr>
<tr>
<td>RNase-free water</td>
<td>8 μl</td>
<td></td>
</tr>
<tr>
<td>Total Reaction Volume</td>
<td>20 μl</td>
<td></td>
</tr>
<tr>
<td>Step</td>
<td>Time</td>
<td>Temperature</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>Initial activation step</strong></td>
<td>15 mins</td>
<td>95°C</td>
</tr>
<tr>
<td><strong>Cycling step</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denaturation</td>
<td>15 sec</td>
<td>95°C</td>
</tr>
<tr>
<td>Annealing</td>
<td>30 sec</td>
<td>64°C</td>
</tr>
<tr>
<td>Extension</td>
<td>30 sec</td>
<td>72°C</td>
</tr>
<tr>
<td><strong>Number of cycles</strong></td>
<td>40</td>
<td></td>
</tr>
</tbody>
</table>

**4.2.17. Restriction Digest Test**

Restriction enzymes (restriction endonucleases) are enzymes that are able to recognize short DNA sequences. They cleave double-stranded DNA (dsDNA) at specific sites recognition sequences. Each restriction enzyme has specific conditions such as temperature, pH, enzyme cofactor(s), salt composition and ionic strength for optimal activity. In this study, restriction enzymes were purchased from Promega. The protocol recommended by manufacturer was followed to verify the orientation of the cloned fragments. In a sterile tube, the following components were assembled in the order listed below in Table 4.16. The reaction mixture was mixed gently by pipetting, centrifuged for a few seconds in a microcentrifuge. Following, the mixture was incubated at the enzyme’s optimum temperature for 1–4 hours.
Table 4.16: Reaction components for restriction digest test.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume of Samples and Standards (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deionised water</td>
<td>16.3</td>
</tr>
<tr>
<td>Restriction Enzyme Buffer 10X</td>
<td>2</td>
</tr>
<tr>
<td>Acetylated BSA (10 μg/μl)</td>
<td>0.2</td>
</tr>
<tr>
<td>DNA, 1 μg/μl</td>
<td>2</td>
</tr>
<tr>
<td>Probe</td>
<td>2</td>
</tr>
<tr>
<td>Components above were mixed by pipetting and then enzyme was added:</td>
<td></td>
</tr>
<tr>
<td>Restriction Enzyme, 10 u/μl</td>
<td>0.5 10 μl</td>
</tr>
</tbody>
</table>

4.2.18. FTIR Spectroscopy

FTIR spectroscopy is a fingerprint method to identify unknown materials, determine the quality of and consistency of samples. The analysis has significant advantages over other imaging methods since it based on detecting the characteristic absorbance of corresponding molecular vibrations in the sample. It requires a small amount of sample with little or no sample preparation prior, which provides a quick analysis of samples. Another key advantage is that the method does not require any additives (dye, label) for visualization.

FTIR spectroscopy allows identifying the functional groups in the sample. For this purpose, FTIR spectroscopy was used to identify the functional group in cellulose samples produced by bacteria. ATR-IR spectra were recorded using a Spectrum One FTIR-spectrometer (Perkin Elmer, Massachusetts, USA). The spectra were collected at a resolution of 2 cm⁻¹ in the range of 600 and 4,000 cm⁻¹.
4.2.19. Scanning Electron Microscopy

Scanning electron microscopy (SEM) visualizes the surface of materials by using electrons instead of light to form an image. The signals that derive from electron-sample interactions give information about external morphology (texture), crystalline structure of the sample and orientation of materials that constitutes the sample. SEM allows producing images of high resolution and represents the structures of samples in three-dimensional. Sample preparation is relatively easy since most SEMs only require sample coating in order to generate conductive samples.

The morphology and the microstructural features of samples were investigated using a field emission scanning electron microscope (FESEM), LEO 1525 (Zeiss, Germany) operating at 5kV. Prior to analysis, samples were dried until constant weight and placed on a double-sided carbon tape mounted onto an aluminium stub. They were gold coated for 2 min at 20mA using an Emitech K575X Peltier (Ashford, UK) cooled sputter coater.

4.2.20. Congo Binding Assay

Congo red binding assay is a method to detect bacterial colonies able to use cellulose. It has been observed that Congo red forms complexes with unhydrolyzed polysaccharides.

Cellulose production was assayed by growing the *E. coli* cells until the OD$_{600}$ reached 1.0 and inducing the cells with 0.2 mM IPTG to initiate the gene expression for the production of BC. After that, the cells were diluted and plated as explained in Chapter 3.2.6 on LB agar plate containing 25 mg/L Congo red, 50 mg/ml streptomycin, 50 mg/ml carbenicillin. As control, the *E. coli* cells grown without IPTG induction also plated on LB agar plate. All plates were incubated at 37 °C for 24 hr and screened qualitatively to detect the BC production ability. Also, *E. coli* BL21/HMS (DE3) (pBCS + pCMP) induced by 0.2 mM IPTG was grown 18 hours after induction. After that the culture supplied with 25 mg/L Congo red and incubated in
static conditions in room temperature and the cells were harvested by centrifugation at 4000 rpm for 10 minutes for qualitative analysis.

### 4.2.21. Quantification of Bacterial Cellulose Production

Bacterial cellulose samples were collected from culture media by filtration. Falcon cell strainers with 40 μm mesh size were used to filter the cultures containing cellulose. Following the separation of cellulose samples from culture by filtration, samples were washed with 70% ethyl alcohol and distilled water several times. After that, samples were left to dry overnight at room temperature. Finally, total sample weight was determined by balance with minimum capacity to detect 0.1 mg.

### 4.2.22. Determination of Free Carbon Source

In this study, glucose and glycerol were used as carbon source for the production of cellulose. In order to detect the consumption during culturing, enzymatic assay were used.

#### 4.2.22.1. Enzymatic Glucose Assay

The principle of the assay based on a chain reaction with two steps. Hexokinase (EC 2.7.1.1) catalyses the formation of glucose-6-phosphate (Glc-6-P) from glucose [109]. In the first step, glucose is phosphorylated by adenosine triphosphate (ATP) in the presence of hexokinase. In the second step, Glc-6-P is oxidized to gluconolactone 6-phosphate by glucose-6-phosphate dehydrogenase (G6PDH) while NADP+ is reduced to NADPH. Hence, Glc-6-P can be determined by measuring NADPH absorbance at 340 nm. During the oxidation, an equimolar amount of NAD is reduced to NADH. The consequent increase in absorbance at 340 nm is directly proportional to the level of glucose in the sample.

\[
\text{Hexokinase} \quad \text{Glucose} + \text{ATP} \rightarrow \text{Glc-6-P} + \text{ADP}
\]

\[
\text{G6PDH} \quad \text{Glc-6-P} + \text{NAD} \rightarrow \text{6-Phosphogluconate} + \text{NADH}
\]
In this study, Glucose Assay Reagent (Sigma) that consists of 1.5 mM NAD, 1.0 mM ATP, 1.0 U/ml of hexokinase, and 1.0 U/ml of glucose-6-phosphate dehydrogenase with sodium benzoate and potassium sorbate as preservatives was used. In order to determine the amount of free glucose in culture medium, the samples were collected every two hours and centrifuged at 10000 rpm for 10 minutes to separate the cells from medium. Following the sterilization of supernatant by filter, they stored for free glucose determination assay. The composition of reaction mixture was summarized in Table 4.17. The absorbance was detected by UV spectrophotometer at 330 nm.

**Table 4.17:** Composition of reaction mixture for glucose determination assay.

<table>
<thead>
<tr>
<th></th>
<th>Volume of Glucose Assay Reagent</th>
<th>Sample Volume</th>
<th>Volume of Deionized Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Blank</td>
<td>--</td>
<td>10 μl</td>
<td>1.0 ml</td>
</tr>
<tr>
<td>Reagent Blanks</td>
<td>1.0 ml</td>
<td>--</td>
<td>10 μl</td>
</tr>
<tr>
<td>Sample</td>
<td>1.0 ml</td>
<td>10 μl</td>
<td>--</td>
</tr>
</tbody>
</table>

In order to calculate the accurate amount of free glucose in medium, following equations were used.

\[
A_{Total \ Blank} = A_{Sample \ Blank} + A_{Reagent \ Blank}
\]

\[
\text{mg glucose/ml} = \frac{(\Delta A) (TV) (\text{Glucose Molecular Weight}) (F)}{(e)(d)(SV)(\text{Conversion Factor for \( \mu g \) to \( mg \)})}
\]

\[
\text{mg glucose/ml} = \frac{(\Delta A) (TV) (180.2)(F)}{(6.22) (1) (SV) (1,000)}
\]

\[
\text{mg glucose/ml} = \frac{(\Delta A) (TV) (0.029)}{(SV)}
\]
\[ \Delta A = A_{\text{Test}} - A_{\text{Total Blank}} \]

TV = Total Assay Volume (ml)
SV = Sample Volume (ml)
Glucose MW = 180.2 g/mole or equivalently 180.2 µg/µmole
F = Dilution Factor from Sample Preparation
\( \varepsilon \) = Millimolar Extinction Coefficient for NADH at 340 nm (ml/µmole)(1/cm)
\( d \) = Light path (cm) = 1 cm
1,000 = Conversion Factor for µg to mg

**4.2.22.2. Enzymatic Glycerol Assay**

The principle of the assay based on a chain reaction with three steps. In the first step of the chain, glycerol is phosphorylated by adenosine triphosphate (ATP) in the presence of glycerol kinase (GK) which catalyses the formation of glycerol-1-phosphate (G-1-P) and adenosine-5'-diphosphate (ADP). In the second step, G-1-P is then oxidized by glycerol phosphate oxidase (GPO) to dihydroxyacetone phosphate (DAP) and hydrogen peroxide (H\(_2\)O\(_2\)). Following this step, peroxidase (POD) catalyses the coupling of H\(_2\)O\(_2\) with 4-aminoantipyrine (4-AAP) and sodium N-ethyl-N-(3-sulfopropyl) m-anisidine (ESPA). This reaction produces a quinoneimine dye that gives a pink colour to reaction mixture. The increase in absorbance at 540 nm is directly proportional to the free glycerol concentration of the sample.

![Enzymatic Glycerol Assay Diagram](image)

In this study, Free Glycerol Reagent (Sigma) was used which consists of 0.75 mM ATP, 3.75 mM magnesium salt, 0.188 mM 4-Aminoantipyrine, 2.11 mM N-Ethyl-N-(3-sulfopropyl) m-anisidine sodium salt, 1,250 U/L glycerol kinase (microbial), 2,500 U/L glycerol phosphate oxidase (microbial), 2,500 U/L Peroxidase
(horseradish), 0.05% sodium azide. In order to determine the amount of free glycerol in culture medium, the samples were collected every two hours and centrifuged at 10000 rpm for 10 minutes to separate the cells from medium. Following the sterilization of supernatant by filter, they stored for free glycerol determination assay.

Table 4.18: Composition of reaction mixture for glycerol determination assay.

<table>
<thead>
<tr>
<th></th>
<th>Volume of Glucose Assay Reagent</th>
<th>Sample Volume</th>
<th>Volume of Deionized Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blank</td>
<td>0.8 ml</td>
<td>--</td>
<td>10 μl</td>
</tr>
<tr>
<td>Standart</td>
<td>0.8 ml</td>
<td>10 μl (0.26 mg/ml)</td>
<td>--</td>
</tr>
<tr>
<td>Sample</td>
<td>0.8 ml</td>
<td>10 μl (1:10 diluted)</td>
<td>--</td>
</tr>
</tbody>
</table>

Blank, Standard, and Sample are prepared according to Table 4.18. Reaction mixtures were mixed by gentle inversion. Following the incubation for 5 minutes at 37 °C, absorbance was recorded at 540 nm versus water as reference. Glycerol concentration of the each sample was calculated by following equation.

\[
glycerol\ content = \frac{A_{sample} - A_{blank}}{A_{standart} - A_{blank}} \times Concentration\ of\ standard
\]

4.2.22.3. Acetate Determination Assay

The principle of the assay based on couple of enzymatic reactions. Acetate is an important component for biosynthesis. Acetate and acetate substrate mix converted to an intermediate in the reaction catalysed by Acetate Enzyme Mix. The intermediate reduces the probe to a yellow-colored product and absorbance was read at 450 nm.
In order to determine the amount of produced acetate in culture medium, the samples were collected every two hours and centrifuged at 10000 rpm for 10 min to separate the cells from medium. Following the sterilization of supernatant by filter, they were stored for acetate determination assay.

0, 2, 4, 6, 8, and 10 mL of the 1 mM standard acetate solution were added into a 96 well plate, generating 0 (blank), 2, 4, 6, 8, and 10 nmole/well standards. After that, acetate assay buffer was added to each well to complete the volume to 50 μl. The reaction mixtures were set up according to the scheme in Table 4.19.

Table 4.19: Composition of reaction mixture for acetate determination assay.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume of Samples and Standards (μl)</th>
<th>Blank (μl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetate Assay Buffer</td>
<td>42</td>
<td>44</td>
</tr>
<tr>
<td>Acetate Enzyme Mix</td>
<td>2</td>
<td>--</td>
</tr>
<tr>
<td>ATP</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Acetate Substrate Mix</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Probe</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

50 μl of reaction mixture was added to each well of 96 well plates. Following the 40 minutes incubation at room temperature, absorbance was measured by machine at 450 nm.
4.2.23. Operation of Stirred-tank Bioreactor

An instrumented stirred-tank bioreactor also used in this study for cultivation of *E. coli* strains. The control unit enabled to control various parameters such as pH, temperature, agitation speed, and O₂ concentration. Tight control over parameters produces the most satisfactory results including maximized cell growth and productivity. The bioreactor was operated at 22°C, 30 °C, and 37 °C. At 22 °C, the chilling unit was continuously maintaining the temperature constant by circulating cold water it in a closed-loop within a cooling-coil inside the bioreactor vessel. At 30 °C and 37 °C, heating jacked was maintaining the temperatures constant. In order to keep the pH at 7.0±0.1, 1 N NAOH and 0.5 M HCl was connected to peristaltic pump A and B, respectively. O₂ concentration was set 20% and air flow provided by a pump to keep the O₂ concentration constant. The agitation speed was set at 250 rpm. The total volume of medium was 4 L.

![Figure 4.7: Instrumented-bioreactor equipped with feedback control-loop[110].](image)

4.2.24. Statistical Analysis

Each fermentation was repeated three independent times and each analysis was replicated at three times. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to P < 0.05.
5. DESIGN OF THE RECOMBINANT EXPRESSION SYSTEM
5.1. Introduction and Aim

Recombinant DNA technology is a very powerful tool, which allows production of recombinant proteins for industrial purposes. Bacterial expression systems for heterologous protein production draw attentions as their ability to grow rapidly and at high density on inexpensive substrates for the maximum production of desired product. Also, their genetics are often well-characterized and large number of cloning vectors and host strains are available.

Cellulose synthase, which regulates the synthesis of cellulose from UDP-glucose, is encoded by four \((bcsA, bcsB, bcsC, \text{ and } bcsD)\) subunits in \(A. xylinum\) ATCC53582. In a recent study, functional cellulose synthase has been tried to reconstructed in \(E. coli\) by simultaneous expression of \(bcsA\) and \(bcsB\), however, reconstruction was incomplete. This was attributed to absence of \(bcsC\) and \(bcsD\) [111]. The upstream region of the \(bcsABCD\) operon has two genes; \(cmcax\) and \(ccpAx\), which play role in ribbon assembly and transport of cellulose chains. The function of \(bcsABCD\) and its upstream operon is explained earlier in Chapter 2.4 in detail. It has been reported that \(bcsABCD\) operon and the upstream operon is essential for the production of cellulose. For this reason it is necessary to clone all these 6 cellulose synthesis related genes to achieve the cellulose biosynthesis in \(E. coli\). In addition, it has been mentioned that the disruption of the \(bglxA\) gene, located downstream region of \(bcsABCD\) operon, causes a decrease in BC production [69].

DNA sequences with the length of 3-10 kbp are considered as large fragments for a cloning, which decreases the efficiency of amplification and cloning significantly. The total length of the target sequence for cellulose biosynthesis in this work is 14.5 kbp, which is far longer than 10 kbp, considered as an extra-long fragment. This is one of the significant bottlenecks in the achievement of BC production in \(E. coli\) by recombinant DNA technology. Selection of a suitable and efficient expression system for only one gene is often considered problematic in literature. Therefore, functional co-expression of the all cellulose synthase related proteins \((bcsABCD, cmcax, ccpAx, bglxA)\) to achieve cellulose production is
extremely complicated. For this reason, selection of the suitable approach among many options is extremely tricky. The work reported in this chapter explains the details of the design of the expression system for the bacterial cellulose biosynthesis. The important factors for successful expression of proteins were summarized in Table 5.1, which are discussed in this chapter.

Table 5.1: Important factors for a successful recombinant expression.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of host</td>
<td>Eukaryotic/prokaryotic, yeast/bacteria.</td>
</tr>
<tr>
<td>Type of plasmid/plasmids</td>
<td></td>
</tr>
<tr>
<td>Number of target proteins</td>
<td>Co-expression/multiple vector/single vector</td>
</tr>
<tr>
<td>Origin of replication</td>
<td>pMB1 ori, ColE1 ori, p15A ori, pUC ori etc.</td>
</tr>
<tr>
<td>Copy number</td>
<td>High copy number or low copy number.</td>
</tr>
<tr>
<td>Selection Marker</td>
<td>Streptomycin, ampicillin, kanamycin etc.</td>
</tr>
<tr>
<td>Promoter</td>
<td>Strong/weak, inducible/constitutive</td>
</tr>
<tr>
<td>Codon usage</td>
<td>Optimised according to host strain</td>
</tr>
<tr>
<td>Temperature</td>
<td>Protein folding and stability, plasmid stability.</td>
</tr>
<tr>
<td>Growth conditions</td>
<td>Oxygen levels, pH, carbon source and fermentation strategy.</td>
</tr>
<tr>
<td>Inducer concentration</td>
<td>Plasmid stability, metabolic stress.</td>
</tr>
</tbody>
</table>

The presented chapter is composed of 10 parts: introduction and aim, methodology, choice of the host cells, choice of the plasmid, design of the expression system, cloning strategy, restriction digest test results, plasmid transformation into *E. coli* cells, discussion and conclusion.
5.2. Methodology

In this chapter, characteristics of a microbial recombinant expression system was evaluated and discussed. Two different approaches were mentioned for the plasmid design: (i) One plasmid approach (ii) Two plasmid approach. For the first approach, DNA isolation and PCR experiments were performed by following the protocols presented in Chapter 4.2.10 and Chapter 4.2.12, respectively. PCR result was verified by agarose gel electroporation. For the second approach, target fragments were synthetically produced, optimised and cloned by Genewiz, Inc. Restriction digest test was performed for each plasmid to verify the orientation of the cloned fragments. Host \textit{E. coli} cells was transformed by recombinant plasmids by the protocol presented in Chapter 4.2.7.

5.3. The Selection of The Host Cell

The selection of the host strain is the first and the most important step of the design of an expression system. The choice of the host cell as protein synthesis machinery for the production of the target proteins initiates the structure of the whole process in terms of the technology, molecular tools, equipment, or reagents needed for the project. There are large numbers of available host systems based on microorganisms such as bacteria, yeast, filamentous fungi, and unicellular algae. All of them have advantages and disadvantages which affects the efficiency of the system [112, 113].

Among microorganisms, the gram-negative bacteria \textit{E. coli} is the most commonly used organism for recombinant protein production. One of the reasons is the fact that \textit{E. coli} is a well-characterised model organism which has been extensively studied on. It can survive in various growth conditions in inexpensive culture media. Large scale production systems which are established on \textit{E. coli} allow a rapid and economical production of desired products as its doubling time is approximately 20 minutes [78]. Also, transformation with plasmid DNA is straightforward, which can be performed in 5 minutes [114].
The other important characteristic of *E. coli* is the fact that some of the pathogenic and commensal strains are able to produce cellulose. Production of cellulose is under the control of *bcsABCZ* operon, which is highly homologous to the cellulose biosynthesis operon in *A. xylinum*. Although existence of those genes in *E. coli* is promising, their existence, function and conditions for production remains elucidated. Also, cellulose production occurs only when *adrA*, a gene encoding a putative transmembrane protein, is expressed. For the high yields of production of cellulose by non-pathogenic strains that can reach high optical densities, there is a need to use genetic modification on industrially valuable strains [115, 116].

There is large number of suitable *E. coli* candidates in literature to use as hosts for over-expression of target proteins. All of them have different characteristics and most of them are used in specific situations. The most widely used *E. coli* strains are BL21 and K-12 derivatives.

*E. coli* BL21 is a B derivative, which is deficient to code two proteins that are not desired in a recombinant expression system. It is not able to express Lon protease, which degrades many foreign proteins [117]. Additionally, it is also deficient in expressing *OmpT*, which degrades extracellular proteins. Deficiency in coding these proteins makes this strain significantly advantageous since that decreases the breakdown of the proteins. Also it produces considerably lower amount of acetate, which caused inhibition in growth when it is accumulated in the growth media, compared to other strains of *E. coli* [118]. In the BL21 (DE3) strain, the λDE3 prophage is inserted in the chromosome of BL21 and contains the T7 RNA polymerase gene, which is under the control of inducable *lacUV5* promoter. This is a strong variant of the wild-type *lac* promoter, which can be simply controlled by IPTG. Some of the derivatives of BL21 (DE3) are designed for the toxic protein expression. C41 (DE3) and C43 (DE3) exhibit a significant improvement in the expression of membrane proteins, as membrane proteins are often toxic [119]. Expression of recombinant proteins were controlled also by *lacUV5* promoter in these strains, as some of the regions revert this promoter back into weaker wild-type counterpart [120]. The other important strain for industrial use is K-12 and its
derivatives. Especially HMS174 is a widely used derivative of K-12 [121]. This strain has a recA mutation, which promotes plasmid stability [122]. This strain has also a λDE3 derivative that allows the use of T7 TNA polymerase system.

Considering the advantages of λDE3 derivatives of BL21 and K-12 strains, plasmids were aimed to design compatible with both derivatives to investigate the cellulose production in different strains.

5.4. The Choice of The Plasmid

There is a wide range of plasmids available today with multiple combinations of origin of replications, promoters, selection markers and multiple cloning sides, as presented in Figure 5.1. For this reason, it is important to evaluate all the major features in an expression vector.

![Figure 5.1: Major features of an expression vector.](image-url)
5.4.1. Major Features of Expression Vectors

Origin of replication (ORI) or replicon is the sequence where plasmid DNA replication begins. This allows a plasmid to reproduce itself inside cells. ORI determines the plasmid copy number, since the control of copy number takes place near the plasmid's origin of replication. ORI also determines the compatibility of two plasmids, as different plasmids replicated by the same origin cannot co-exist in the same cell. Copy numbers and compatibility groups were summarized in Table 5.2. Plasmids in the same incompatibility group cannot co-transformed.

One of the most important parameter for the selection of suitable vector is copy number [123]. The use of high copy number plasmids often results in high metabolic stress on bacteria, which limits cell growth[76]. Especially expression of the membrane proteins, which are often toxic, by high copy number plasmids could cause the accumulation of proteins in high levels since the presence of hundreds of copies of plasmids. For this reason, the use of high copy number plasmids could not guarantee a high yield production. A low copy number plasmid could reduce the toxic level of expression.

<table>
<thead>
<tr>
<th>Origin of Replication</th>
<th>Copy Number</th>
<th>Incompatibility Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMB1</td>
<td>~15-20</td>
<td>A</td>
</tr>
<tr>
<td>Modified pMB1</td>
<td>~500-700</td>
<td>A</td>
</tr>
<tr>
<td>ColE1</td>
<td>~40</td>
<td>A</td>
</tr>
<tr>
<td>p15A</td>
<td>~18-22</td>
<td>B</td>
</tr>
<tr>
<td>pSC101</td>
<td>~5</td>
<td>B</td>
</tr>
<tr>
<td>CloDF13</td>
<td>~20-40</td>
<td>B</td>
</tr>
</tbody>
</table>

In order to initiate the transcription, a promoter is required in vectors, which located near the start codon of target gene. The promoters used in recombinant expression vectors are generally inducible with the introduction of an
inducer, which allows control of the expression. Protein expression could achieve also by constitutive promoter. However constant expression of proteins may cause extra metabolic stress especially in the lag phase of the growth.

The most widely used promoter is lac promoter in a recombinant expression vector, which can be induced by lactose. However, induction is not always achieved when the inducer presented into media, because efficient transcription occurs in the presence of sufficient concentration of cyclic AMP (cAMP) and cyclic AMP protein (CAP). The CAP/cAMP complex, which binds to just upstream of the lac promoter, stimulates transcription by RNA polymerase. The concentration of cAMP is directly dependent on carbon source. Also, glucose inhibits lactose uptake since lactose permease is not active in the presence of glucose [124]. In the presence of high concentrations of glucose, the concentration of cAMP is low and this phenomenon is called catabolite repression [125]. Therefore the full induction of the lac operon is accomplished in the presence of inducer if the cAMP levels are sufficient. A mutant of lac promoter, lacUV5, is developed to overcome the effect of catabolic repression. However when the copy number of the plasmid is high, a leak of expression can be detected in high levels in the absence of inducer due to low level of expression of the lac promoter repression [126]. lac promoter and its derivative lacUV5 are not very useful for recombinant expression due to their weaknesses mentioned above.

The T7lac promoter system, on the other hand, has significant advantages over lac promoter. The plasmids developed based on this system carry T7 promoter, lac operator and rbs. In T7lac expression system, target gene is cloned downstream of a promoter, which recognised by T7 RNA polymerase (T7 RNAP). This enzyme usually is placed in bacterial genome in a prophage (λDE3), under the control of lacUV5 promoter [125, 127]. When the inducer introduced into media, it initiates the expression of T7 RNAP. Following that, T7 RNAP binds to T7 promoter and this initiates the expression of target gene. In order to supress the basal expression, plasmids with T7lac promoter has lac repressor, which encodes a
protein binding to lac operator, in the absence of inducer. Target protein can represent 50% of the total cell protein in this system [128, 129].

In E. coli expression systems, a selective pressure must be imposed for stability of the plasmid during expression. Antibiotic resistance genes are often used as selection markers to ensure the maintenance of plasmid inside cells. Thus, addition of appropriate antibiotic to the medium would force cells to maintain the plasmid carrying antibiotic resistance in order to survive. Ampicillin is one of the most popular antibiotics used in expression vectors. This gene encodes β-lactamase, which cleaves β-lactam ring of β-lactam antibiotics. However, ampicillin is not a durable antibiotic and depleted in a couple of hours [130]. Chloramphenicol is another antibiotic used in recombinant expression, which inhibits blocks the peptidyltransferase reaction. Cam\textsuperscript{R} gene provided in plasmids encodes cytoplasmic chloramphenicol acyltransferase, which inactivates chloramphenicol by covalently acetylation it. Streptomycin inhibits protein synthesis by binding to the S12 protein. Str\textsuperscript{R} gene prevents streptomycin from binding to the ribosome.

Variations in codon usage between different bacteria also have a significant influence on recombinant protein expression. This phenomenon is called codon bias, which occurs when the codon usage of the host organism differs from the original source of the protein. When the foreign mRNA contains many codons that are rare in E. coli, the translation rate could decrease with a low level of their tRNAs species in the cell. Moreover, codon bias can cause a decrease in mRNA stability, premature termination of translation, frameshifts and inhibition of protein synthesis and cell growth. Consequently, low expression levels could be observed. Approaches for solving codon usage bias such as codon optimization of the foreign coding sequence by silent mutagenesis or increasing the availability of tRNAs by host modification have been described in literature [131-133]. Significant increases have been achieved in expression levels by both approaches [131]. Analysis showed that overall GC content is the most important parameter that influences the differentiation in codon bias [134]. A very low or very high GC composition is related with a large codon usage bias. Genes from species with similar phylogeny or
with similar tRNA content have similar codon bias and an organism’s optimal
growth temperature influences the codon bias of its genes [135, 136]. A high GC
content in the 5'-end of the gene of interest usually results in the formation of
secondary structure in the mRNA. This could lead to interrupted translation and
lower levels of expression. Thus, higher expression levels could be obtained by
changing G and C residues at the 5'-end of the coding sequence to A and T residues
without changing the amino acids.

5.4.2. Protein Co-expression and Duet Vectors

Protein co-expression could be accomplished by multiple or single plasmids.
The easiest approach for co-expression is to design a system using vectors with
different selection makers and to express each gene by a single vector. Each
plasmid requires to carry different selection markers and to have different
compatible replications. Although it has been published that the plasmids with the
same origin of replication can be used as long as they carry different selection
markers and this could cause a competition inside the cells. As a result, one of the
plasmids could be replicated with poor stoichiometry. To use vectors with different
origins of replications has been proved a widely used approach for protein co-
expression [137-139].

It is also possible to insert several genes into the same vector, when the
vector has multiple cloning sides. Transcription is initiated by either individual
promoters or a single promoter. When all genes cloned under the control of a
single promoter, each gene has its own ribosome binding side, which results in a
long polycistronic mRNA. In this case, the expression level of the proteins depends
on the order of the proteins on a plasmid, especially for the long sequences, since
the stability of mRNA is usually problematic. It has been proved that polycistronic
expression causes a lower expression of the more downstream encoded protein.
This could negatively affect the stoichiometry of a co-expression. On the other
hand, use of individual promoters resulted in several fold higher expression than
polycistronic expression. Although the size of the plasmid is bigger when individual promoters are used, mRNAs are smaller [140, 141].

A novel co-expression system was developed by Novagen based on Duet vectors, which presents many advantages in co-expression of recombinant proteins. Duet vectors makes possible to express 8 different genes in 4 different plasmids, having difference selection markers and compatible origin of replications [142, 143]. The compatible replicons and antibiotic resistance genes are designed to achieve an effective propagation and maintenance of four different plasmids in a single *E. coli* cell. Use of vectors carrying different resistance markers and compatible origin of replications has been exhibited successful expressions [144, 145]. Each of these vectors carries two expression units, each controlled by a separate T7lac promoter, which provides a tight control of the expression by IPTG induction. Moreover, use of a single promoter for each gene results in higher yields compared to polycistronic expression. Each expression unit has own optimal ribosomal binding side and multiple cloning side (MCS) located downstream of the promoter. In each vector, MCS1 begins with an Nco I site (CCATGG) at the ATG (Met) translation initiation site, for the expression of unfused proteins. This is followed by a His-tag and various restriction sites, which are used commonly such as BamH I, EcoR I, Sac I, Sal I, Hind III. MCS2 begins with Nde I (CATATG) site at the ATG (Met) translation initiation site for also the production of unfused protein. MCS2 harbours Bgl II, Mun I, and Xho I sites to generate overhangs compatible with BamH I, EcoR I, and Sal I overhangs, respectively. MCS2 also harbors S•Tag™ peptide in downstream.

Separate expression units allow keeping the stoichiometry of a co-expression with a high yield expression. Relative copy number of all Duet plasmids are pRSFDuet-11 >pETDuet-1 > pCDFDuet-1 > pACYCDuet-1, as presented in Table 5.3. They are also compatible a set of *E. coli* strains for co-expression.
### Table 5.3: Origin of replications and copy numbers of Duet vectors.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Origin of Replication</th>
<th>Copy number</th>
</tr>
</thead>
<tbody>
<tr>
<td>pETDuet-1</td>
<td>ColE1</td>
<td>~40</td>
</tr>
<tr>
<td>pACYCDuet-1</td>
<td>p15A</td>
<td>10-12</td>
</tr>
<tr>
<td>pCDFDuet-1</td>
<td>CloDF13</td>
<td>20-40</td>
</tr>
<tr>
<td>pRSFDuet-1</td>
<td>RSF1030</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

pETDuet-1 has the ColE1 replicon and bla gene for ampicillin resistance, pACYCDuet-1 has the P15A replicon and cat gene for chloramphenicol resistance, and pCDFDuet-1 has the CloDF13 replicon and aadA gene for streptomycin resistance, and pRSFDuet-1 has the RSF1030 replicon and kanamycin resistance. As they all have different origin of replications and antibiotic resistances, pETDuet-1 vector can be transformed into the same cell with pCDFDuet-1, pACYCDuet-1, and/or pRSFDuet-1 in the compatible host strain for the co-expression of up to 8 target genes. However, host strain compatibility should be evaluated. Novagen vectors have been used in large number of studies efficiently for co-expression [146-149].

The nature and positioning of the cloning sites in the Duet vectors generate several options for expression, detection, purification, and quantification of protein complexes. As a consequence of evaluation of characteristics and advantages of the Duet vectors, the recombinant expression system designed for the production of bacterial cellulose aimed to derive from Duet Vectors.
5.5. Design of the Expression System

5.5.1. One Plasmid Approach

Cloning bcsABCD operon, its upstream and downstream operons into a pET derived vector was evaluated as a first approach in an effort to design a straightforward system to control. The schematic representation of the design is presented in Figure 5.2.

According to this design, all of the genes were under the control of the same promoter. However, this was extremely problematic, as the target fragment is an extra-long for a cloning (14.5 kbp). Although PCR is one of the most essential tools in molecular biology to clone small DNA fragments, the application of PCR reactions were limited by the size of amplified fragments. Invention of long-range PCR increased the size of amplicons from 3–5 kbp to over 30 kbp by modifying the polymerases [150-152]. However, this is possible under highly optimised conditions, which depend on many parameters: concentration of the components (DNA polymerase, primers, DNA, magnesium), primer design, annealing temperature, annealing time, extension time, purity of DNA sample, etc. As a consequence, amplification of the target sequence by long range polymerase chain reaction (PCR) was very inefficient. In this study, genomic DNA of A. xylinum was isolated and amplified by PCR. The total target fragment was only in %2 of the total number of PCR successfully amplified, which was not sufficient in terms of volumetric yield of the fragment to continue with cloning experiments. Also, when the PCR product
was visualised by agarose gel electrophoresis for the confirmation of the amplification, as presented in Figure 5.3, the DNA sample exhibited a weak and smeared band. This could occur as a result of many reasons such as non-optimal concentrations of Mg$^{2+}$, nucleotide or DNA, non-optimal annealing temperature, or impurities in sample. The purity of the sample is one of the fundamental requirements for cloning.

**Figure 5.3:** Amplification of target fragment (bcsABCD, cmcAUX, ccpAX). Lane 1: molecular weight marker, Lane 2: PCR product.

In addition to the inefficiency of the PCR reaction conducted in our laboratory conditions, other disadvantages in this approach were also evaluated. As mentioned in this chapter earlier, the expression level of the proteins depends on the order of the proteins on a plasmid, as a consequence, the proteins located the downstream of the plasmid could be expressed in a low level. Moreover, the target fragment is an extra-long for a cloning, which could results in a plasmid with the size of bigger than 15 kbp. The size of a plasmid directly influences the plasmid stability, as expression systems contain bigger plasmids more likely to have plasmid stability problem driven by metabolic stress [153, 154]. Also transformation efficiency decreases with increases of size of the plasmid [155]. Another weakness of this system was detected by the sequence analyses performed by The ORF Finder (Open Reading Frame Finder), which is an analysis tool which finds all open reading frames of a sequence. Genes that code for proteins include open reading frames
(ORFs) consisting of a series of codons that specify the amino acid sequence of the protein, which starts with a start codon and ends with a stop codon. Each DNA sequence has six reading frames, three in one direction and three in the reverse direction on the complementary strand, as presented in Figure 5.4. bcsA and bcsC were detected in ORF (+1), while bcsB and bcsD were detected in ORF (+2) by ORF Finder. In order to control the expression of the genes inserted into a cloning vector, the genes should be in the same open reading frame with the promoter. Thus, a polycistronic message could be initiated by the promoter [156]. Otherwise, the genes is oriented in different ORFs could not be under the tight control of the promoter. The sequence properties of the target fragment did not allow the control of the expression by the one single promoter as proposed in the design. Justifying all these disadvantages, use of two compatible plasmids were considered as the second option for the design of the expression system.

![Figure 5.4: Open reading frames in a double-stranded DNA.](image)

**5.5.2. Two Plasmid Approach**

In the second approach, recombinant expression system for the production of BC developed by using backbones of two plasmids: petDuet-1 and pCDFDuet-1. Evaluation of characteristics and properties of the Duet Vectors (Novagen) are discussed previously, in Section 5.1.2.2. Originally, Duet vectors allow expression dual expression of two genes in the same plasmid, as each vector has two multiple cloning sides, as presented in Figure 5.5 and Figure 5.6. This could allow initiation the expression of each gene simultaneously by separate T7lac promoter, which provides a tight control of the system. On the other hand, the expression level of
the proteins depends on the order of the genes on a plasmid, especially for the long sequences, since the stability of mRNA is usually problematic. It has been reported that polycistronic expression causes a lower expression of the more downstream encoded protein. Providing separate promoters for each gene could eliminate this problem by transcription of each gene in a distinct mRNA.

Significant decreases in plasmid stability have been reported with increases in the size of a plasmid [157-159]. For this reason system was simplified in the second approach. The downstream operon, which carries bgIXA, was excluded, as it is effective on the amount of production; however, not essential for the biosynthesis. The other six genes are essential for bacterial cellulose synthesis (bcsABCD operon, cmcaa and ccpAx), as mentioned in Chapter 2.4. These genes could be expressed in 3 compatible Duet vectors; however, this would cause a very high metabolic stress for the cells [160, 161]. Also, since the genes in bcsABCD operon facilitate the biosynthesis of cellulose by encoding cellulose synthase enzyme, separating these genes into different plasmids could affect the stoichiometry of the expression negatively, as each vector could reproduce different number of copies related with its size, origin of replication, or antibiotic stability, etc. On the other hand, upstream operon could be cloned into another plasmid since ccpAx encodes cellulose complementary protein, whereas cmcaa encodes endo-1,4-glucanase. This could allow decreasing the size of the main plasmid that carries bcsABCD operon and could help to eliminate the plasmid instability risk. In addition, increase in the number of genes expressed by a plasmid could negatively affect the level of expression since the time necessary for the replication depends on the size of the plasmid.
Figure 5.5: Original vector map and cloning regions of pETDuet-1.
Figure 5.6: Original vector map and cloning regions of pCDFDuet-1.
Table 5.4: Sequence landmarks of recombinant expression vectors.

<table>
<thead>
<tr>
<th>Sequence Features</th>
<th>pETDuet-1</th>
<th>pCDFDuet-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7 promoter-1</td>
<td>5404-5420</td>
<td>3765-3781</td>
</tr>
<tr>
<td>Multiple Cloning Side-1 (MCS-1)</td>
<td>69-168</td>
<td>69-168</td>
</tr>
<tr>
<td>T7 promoter-2</td>
<td>214-230</td>
<td>214-230</td>
</tr>
<tr>
<td>Multiple Cloning Side-2 (MCS-2)</td>
<td>297-438</td>
<td>297-438</td>
</tr>
<tr>
<td>T7 terminator</td>
<td>462-509</td>
<td>462-509</td>
</tr>
<tr>
<td>ColE1 origin (ori)</td>
<td>2737-2794</td>
<td>-</td>
</tr>
<tr>
<td>CloDF13 ori</td>
<td>-</td>
<td>1611-2349</td>
</tr>
<tr>
<td>Ampicillin resistance gene (AmpR)</td>
<td>1119-1976</td>
<td>-</td>
</tr>
<tr>
<td>Streptomycin resistance gene (SmR)</td>
<td>-</td>
<td>683-1471</td>
</tr>
<tr>
<td>lacI coding sequence</td>
<td>3931-5013</td>
<td>2562-3641</td>
</tr>
<tr>
<td>F1 origin</td>
<td>567-994</td>
<td>-</td>
</tr>
</tbody>
</table>

Sequence landmarks of each plasmid vector are also shown in Table 5.4 and full sequences of both vectors are presented in Appendix 3. Each of the cloning sides is lead by T7 promoter, lac operator and a ribosome binding side. The replicons of pET Duet-1 and pCDFDuet-1 are ColE1 and CloDF13, respectively, which are compatible with each other. ColE1 replicon has a copy number of approximately 40, whereas CloDF13 replicon has a copy number of between 20 and 40. There is a slight difference in plasmid copy number, which could affect the relative expression of the target proteins. Plasmid copy number is often in a negative correlation with plasmid size, since the time necessary for the replication depends on the size of the plasmid. Longer fragments should be cloned into higher copy number plasmids in order to balance the expression in a co-expression system. For this reason, \( bcsABC \)D operon, with the length of 9.2 kbp, was planned to clone into pETDuet-1, while the upstream operon with the length of 2.1 kbp was aimed to clone into pCDFDuet-1. However the Duet vectors have two multiple cloning sides and four multiple cloning sites were required for pETDuet-1 vector in order to express \( bcsABC \)D operon.
containing four different genes (bcsA, bcsB, bcsC and bcsD). Therefore, two genes were aimed to clone in each multiple cloning site of pETDuet-1. bcsA and bcsB were planned to clone into MCS1, while bcsC and bcsD were aimed to clone into MCS2. T7lac promoter, lac operator and ribosomal binding side already existed in the vectors for the expression of first gene. However, generation of these parts for the extra genes (bcsB and bcsD) cloned into each MCS was required to maintain the stoichiometry. Accordingly, following sequences was aimed to synthesize for each MCS in pETDuet-1: (i) bcsA + T7-lac-rbs + bcsB, (ii) bcsC+ T7-lac-rbs + bcsD. The expression system derived from pETDuet-1 and pCDFduet-1 in this study renamed pBCS and pCMP, respectively. The schematic representation of the MCSs of each vector is presented in Figure 5.7.

This design requires cloning of each gene separately as well as the addition of T7-lac-rbs sequence between bcsA-bcsB and bcsC-bcsD. Construction of pBCS is not as straightforward as the construction of the plasmid design in one plasmid approach. If multiple inserts are present in a single vector, the determination of restriction sides becomes extremely complex. Also, ligation of the inserts in the correct orientation during the cloning processes is extremely difficult, which could
require multi step cloning since there is a need to use an entry vector first. This process also takes time as well as associated with costs. Multi-step cloning could result in undesired products as due to self-ligation. In order to eliminate the disadvantages of multi-cloning process, artificial DNA synthesis was proposed as a solution.

Artificial DNA has been often used for the construction of vectors to expression of recombinant proteins recently, since sequence information from genome projects has increased exponentially over the last decade [162]. This allows application of various modifications on gene sequence to optimise the protein production. Synthetic genes are frequently designed by mimic of natural gene characteristics. Replacing codons that are rarely found in highly expressed E. coli genes with more favourable codons throughout the whole gene has been promoted the expression [163, 164]. It was also suggested that synonymous codon usage beyond the initiation region can impact expression [131, 165, 166]. E. coli strains overexpressing rare tRNAs can significantly improve gene expression [167].

In an effort to enhance the efficiency and accuracy of cloning, while meeting the requirements of the constructs with the correct orientation and codon optimisation, the plasmids have been constructed by Genewiz, Inc, by artificial gene synthesis based on our design.

5.6. Cloning strategy

The cloning procedure was conducted by restriction digest procedure, which is an enzymatic technique. Restriction enzymes cleave plasmid DNA at specific restriction sides, which creates overhangs of 1-6 of nucleotides. In order to be able to clone the target fragments into a vector, both constructs have to carry compatible overhangs for ligation. Generation of compatible overhangs could be achieved by treating both vector and the target fragment with two restriction enzymes that creates compatible ends. This could be done in amplification step (PCR), by designing primers with added compatible sequence of overhangs.
The first step of the cloning procedure based on restriction digestion is the selection of restriction sides of the vector in desired orientation. Different restriction enzymes should be selected for both ends of the insert to achieve the cloning in the correct orientation in order to avoid the self-ligation of the vector or the target fragment. In order to avoid the multiple cuts, it has to be ensured that the selected enzyme does not cleave the target fragment, and there are no repeats of the same restriction side in the vector. In this study, the compatible ends were also synthesized during the synthesis of the target fragments.

For the selection of restriction enzymes, the restriction sides of the multiple cloning sides in each vector were listed. Following that, the sequences of target fragments was run in NEBcutter V2.0, which is a tool detecting cutter (one, two or three cutters) or non-cutter (zero cutters) enzymes in a sequence. When a DNA sequence is run in NEBcutter, a comprehensive report of the restriction enzymes that will cleave the sequence is produced as well as open reading frames are predicted. It gives a diverse output including restriction enzyme maps, theoretical digests and links into the restriction enzyme database. Several options exist to select the enzymes used for digestion, including commercially available enzymes or enzymes that produce compatible termini [168].

As an example, the restriction enzymes that cut the sequence for cmcax just once and the list of enzymes that do not cut the sequence are determined by NEBcutter, presented in Figure 5.8 for cmcax. The lists of zero-cutter enzymes for cmppAx, bcsAB and bcsCD are also provided in Appendix 4.
Figure 5.8: Restriction sites (A) and zero-cutter enzymes (B) for **cmcax**.
Enzymes that are not able to cut the target fragments but able to cleave the vectors were selected for the cloning procedure. The original sequences for *cmcax*, *ccpAx*, bcsA, bcsB, bcsC and bcsD of *A. xylinum* ATCC53582 were obtained from literature. Pubmed access numbers for upstream region (*cmcax*, *ccpAx*) and bcsABCD operon are AB091058 and X54676, respectively [32, 69]. After the selection of restriction sides, the target sequences were analysed by SnapGene, which is molecular biology software developed by GSL Biotech to plan, design and visualise molecular biology procedures. A diverse library of annotated sequence files are provided in the SnapGene web-site including a plasmid library database in proprietary format [169].

The following steps were followed for the cloning strategy for pCMP and analysed by SnapGene:

a. Following the gene optimisation for *E. coli* on the original gene sequence of *ccpAx* obtained from *A. xylinum* ATCC53582, the gene was synthesised including added 5′ (Bgl II) and 3′ (Xho I) sequences. After sequence verification, the gene was cloned into pCDFDuet-1 via 5′ using Bgl II and 3′ Xho I as ligation sides, making the intermediate construct *ccpAx_pCDFDuet-1* for next step.

b. Following the gene optimisation for *E. coli* on original sequence of *cmcax*, the gene was synthesized including 5′ (BamH I) and 3′ (Hind III) sequences. After sequence verification *cmcax* was cloned into *ccpAx_pCDFDuet-1* via 5′ (BamH I) and 3′ (Hind III) as ligation sides to make the final construct *cmcax_ccpAx_pCDFDuet-1*. Schematic representation of the cloning procedure and the circular plasmid map was generated using the SnapGene viewer and presented in Figure 5.9.
The following steps were followed for the cloning strategy for pBCS and analysed by SnapGene:

a. After optimisation of the sequence of bcsA + T7-lac-rbs + bcsB for *E. coli*, the sequence was synthesized including added 5' (BamH I) and 3' (Afl II). After the sequence was verified, it was cloned into pETDuet-1 via 5' BamH I and 3' Afl II to generate the intermediate construct bcsA_bcsB_pETDuet-1 for next step.

b. After optimisation of the sequence of bcsC + T7-lac-rbs + bcsD for *E. coli*, the sequence was synthesized including added 5' (Fse I) and 3' (Xho I). After the sequence was verified, it was cloned into bcsA_bcsB_pETDuet-1 via 5' Fse I and 3' Xho I to generate the final construct bcsA_bcsB_bcsC_bcsD_pETDuet-1. Schematic representation of the cloning procedure and the circular plasmid map was generated using the SnapGene viewer and presented in Figure 5.10.

Full sequence details and GC contents of the optimised fragments (*cmcax, ccpAx, bcsAB* and *bcsCD*) are presented in Appendix 5. Alignments of original and optimised sequences for the genes are presented in Appendix 6.
Figure 5.9: Cloning strategy for pCMP.
Figure 5.10: Cloning strategy for pBCS.
5.7. Restriction Digest Test Results

In order to verify that the inserts were cloned in the correct orientations, restriction digest tests were performed for both plasmids by following protocol explained in Chapter 4.2.17. BamHI and HindIII were used for the restriction of pCMP, which cleaves the plasmid at two edges of the sequences encoding \textit{cmcax}. The reaction was performed at 37°C, which is the optimum temperature for both enzymes. It was expected to result in two linear DNA sequences with the length of 1042 bp (\textit{cmcax}) and 4819 bp (between BamHI and HindIII), respectively. After the restriction procedure, the fragments were visualized by agarose gel analysis, which is presented in Figure 5.11 (A), which clearly showed the DNA bands corresponded to expected sizes in digested sample: 1042 bp and 4819 bp. In addition, unrestricted plasmid was also visualised to verify the total size, which was also identified as a single band corresponding to the total size of pCMP: 5898 bp. A smear was detected in the band of unrestricted sample, which usually arises when the construct is a circular DNA sample, instead of a linear structure. The restriction sides can be seen on pCMP map, in Figure 5.9.

The same procedure was followed to verify the orientation of the inserted fragments in pBCS. Restriction digest test was performed by using MluI and XhoI. Reaction was performed at 37°C, which is the optimum temperature for both enzymes. MluI cleaves the pBCS from two sides, which are located on the sequences of \textit{bcsC} and \textit{lacI} genes, while XhoI cleaves it at the end of the sequences encoding \textit{bcsD}. This was expected to result in three linear DNA sequences with the length of 3071 bp, 4304 bp, and 7219 bp. The DNA bands were detected on agarose gel corresponded to sizes of the fragments that were generated by digestion, as presented in Figure 5.11 (B). The undigested pBCS was also visualised on the gel, which resulted in a smear as a result of circular structure. The restriction sides for pBCS can be found on map, in Figure 5.10.
5.8. Plasmid Transformation into \textit{E. coli} Cells

The transformation of \textit{E. coli} cells by pCMP and pBCS was performed simultaneously. The chemical transformation procedure was followed, which was explained in detail in Chapter 4.2.7. Since the transformation efficiency decreases as the plasmid size increases, the concentration of pBCS was 2.5 times higher than the concentration of pCMP (based on size ratio) in transformation mixture. Selection for transformed cells was accomplished by plating the cells on LB agar containing antibiotics (25 mg/ml carbenicillin and 25 mg/ml streptomycin).

5.9. Discussion

Molecular cloning is a very useful approach for manipulation of the host cells for the production of desired product. A growing number of expression vectors, regulatory element sequences (such as promoters and terminators) and
host cells become available. Despite this progress, many challenges still lie ahead for the co-expression of multiple genes. Consequently, it is essential to select the ideal system for the complex projects.

One of the most significant bottlenecks in the cellulose production by is slow growth of *A. xylinum*. On the other hand, *E. coli* has a significantly fast growth, which allows obtaining large quantities of proteins at low cost in a short time. It remains the most widely used host for prokaryotic expression systems [145, 170, 171].

Cellulose production has been reported in pathogenic and commensal stains of *E. coli* in microscopic levels. The regulator genes of the BC production in *E. coli* have been reported highly homologous to the *bcs* operon of *A. xylinum*. This similarity in the sequences of genes takes part in BC biosynthesis could allow *E. coli* to recognise the cloned fragments. Consequently, *E. coli* selected as the host organism in this study for the expression system. In addition, integration of both DNA-cloning and protein expression technologies in a well-established and high-throughput host like *E. coli* allow parallel testing of different strains and/or culture conditions [172]. As B and K-12 derivatives exhibited many advantages as mentioned in this chapter earlier, Duet Vectors were selected as they are compatible with many B and K-12 derivatives.

Co-expression could be performed by either single or multiple vectors. Both methods were considered in this study. In single plasmid approach, amplification of the total sequence was problematic. The amplification of the target required highly optimised conditions since it was extra-long with a length of 14.5 kbp. Optimisation of the PCR conditions resulted in a highly time-consuming process with high-cost. Also, agarose gel results showed that the samples were not highly pure as a result of multiple binding of primer, which usually occurs in long targets. Purity is one of the most important requirements in complex cloning projects. It has been reported that polycistronic expression causes a lower expression of the more downstream encoded proteins. In order to eliminate this, each gene could be controlled by a separate promoter. However, this requires a multiple cloning process which
includes at least 6 cloning steps and also requires additional sequences for promoter for each target gene (\textit{bcsA}, \textit{bcsB}, \textit{bcsC}, \textit{cmcax} and \textit{ccpAx}). To achieve this complex procedure and clone each gene and its promoter in the correct orientation is highly complex, tricky and unpredictable. On the other hand, use of artificial DNA could be a potential solution. This could give a freedom to organise the promoters and the genes in correct and desired orientation. Furthermore, solving codon usage bias by codon optimization of the foreign coding sequence is highly required for a successful translation of protein encoded by recombinant gene. Also GC content is the most important parameter that influences the differentiation in codon bias affects, which could be optimised by artificial gene synthesis. However artificial synthesis of the target fragments could not be performed in our laboratory conditions. For this reason after the plasmids were design by molecular biology tools (NebCutter, ORF reader, SnapGene), target fragments were produced and cloned by Genewiz, Inc.

When two or more constructs are co-transformed into a single cell, each vector should comprise a different antibiotic selection marker [171, 173] and should harbour incompatible replicons for a stable expression system [139, 171]. One popular approach to express multiple gene expression is to use the Novagen Duet vectors, which carry compatible replication origins (CoLE, CloDF13, P15A, CoLA, and RSF1030) different antibiotic markers (Am, Sm, Cm, Km). They have a versatile plasmid copy number and broad host spectrum. This allows for effective reproduce and maintenance of four plasmids and simultaneous expression of up to eight genes in a single cell [174]. The Duet vectors have been successfully applied to a number of pathways for synthesizing hydroxybutyrate, biofuels, phenylpropanoids, isoprenoids, and polyketides[175-180]. In order to allow their applicability in this study, Duet Vectors were used as a backbone and engineered for the biosynthesis of BC. The backbone of pETDuet-1 was modified with the addition of two more T7\textit{lac} operon and ribosomal binding side to the beginning of the sequences encoding \textit{bcsB} and \textit{bcsD}, in effort to provide separate mRNA translation for each gene. Four genes in \textit{bcsABCD} operon was cloned into pETDuet-1 and the vector renamed pBCS, whereas two genes of upstream operon was
cloned into was cloned into pCDFDuet-1 and the vector renamed pCMP. A very similar approach has been used and succeeded to express six enzymes in the pathway for high-level production of porphyrins in *E. coli* by a system based on pETDuet-1 and pCDFDuet-1. However repeated use of promoter may result in plasmid instability. This could be also a potential risk for this study. Furthermore an increase in the size of an expression plasmid causes an extra metabolic stress on host cells, which could result in plasmid instability. Also the stoichiometry of the expression levels of genes located in the same plasmid could exhibit unbalanced results when the plasmid size increases. In order to eliminate these risks, instead of using one plasmid harbouring all genes, two plasmid approaches was selected. Moreover, this would allow monitoring the effect of *bcsABCD* operon and upstream operon BC production, separately.

The orientations of cloned fragments were verified by restriction digest test and *E. coli* BL21 (DE3) strain, which is a potential high-throughput workhorse, was transformed successfully by chemical transformation as the final step of creation of the expression system.

### 5.10. Conclusion

In this chapter, important factors to achieve recombinant co-expression of multiple genes were justified and ideal approach was selected. Target fragments were cloned and transformed into *E. coli* host strains based on this approach. The next step for the success for an expression system is optimisation of bioprocess parameters, which have significant influence on productivity, such as growth temperature, induction time, inducer concentration, carbon source and culturing strategy. Effect of these parameters on the expression system was studied in following chapters.
6. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED *E. coli* STRAIN 1: BL21 (DE3)
6.1. Introduction

6.1.1. *E. coli* BL21 (DE3) as a Host Microorganism

*E. coli* has a long history of being able to produce many different types of recombinant proteins for several decades since it has many advantages as a host organism. First, it has remarkably fast growth kinetics. Its doubling time is around 20 minutes in glucose-salts media, which means a culture inoculated with a 1/100 dilution of a starter culture can reach stationary phase in a few hours. [78]. However, the expression of a recombinant protein generally imparts metabolic burden on the microorganism, causing a considerable decrease in generation time [76]. As a result of this, high cell density cultures can be easily achieved. The theoretical limit density of an *E. coli* liquid culture is calculated to be around 200 g dry cell weight/l or roughly $1 \times 10^{13}$ viable bacteria/ml [77, 79, 181]. However, batch cultivation of *E. coli* at 37°C, using LB media reaches a cell density of $1 \times 10^{10}$ cells/ml [78], which is less than 0.1% of the theoretical limit. For this reason, high cell-density culture methods were designed to boost *E. coli* growth, even when producing a recombinant protein [182]. Another important advantage is that a rich complex media for the fermentation of *E. coli* has inexpensive components. Additionally, transformation with plasmid DNA is straightforward, which can be performed in 5 min [114].

The B laboratory strains of *E. coli* and its derivatives have been widely used and have had significant impact on basic medicine and biotechnology. the *E. coli* strain B was named by Delbrück and Luria in 1942. It was selected by the phage group that developed around Delbrück, Luria, and Hershey in the 1940s as the host for their studies on the virulent phages T1–T7 [183].

The strain BL21 was derived by various modifications of the B line [184]. BL21 cells are deficient in the Lon protease like as it is in parental B strains. This enzyme degrades many foreign proteins [117]. They are also deficient in a protein coding for the outer membrane protease OmpT, which degrades extracellular proteins. In addition, the level of plasmid loss is low driven by the hsdSB mutation,
which already present in the parental strain (B834). This disrupts the DNA methylation and degradation. In the BL21 (DE3) strain, the λDE3 prophage was inserted in the chromosome of BL21 and contains the T7 RNA polymerase gene under the lacUV5 promoter.

The potential to overexpress recombinant proteins of various *E. coli* strains with different genotypes have been examined. *E. coli* B and *E. coli* K12 strains are almost equally used as host for recombinant protein production (47% and 53%, respectively). However, *E. coli* BL21 is by far the most commonly used strain (35%) in academia. The first reason for this is that BL21 strains are deficient in the production of proteases Lon and OmpT, as mentioned above, which decreases the breakdown of recombinant protein and results in higher yields [117, 185]. Second, *E. coli* BL21 produces a substantially lower amount of acetate compared with K12-derived strains [118].

### 6.1.2. Effects of Plasmid Presence on *E. coli*

Plasmid DNA maintenance and replication commonly induces a “metabolic burden” in *E. coli*. The term “metabolic burden” is defined as the amount of resources (in terms of raw material and energy) that are taken from the host cell metabolism for foreign DNA maintenance and replication [76]. The response of cells under this energy limiting condition caused by metabolic burden is extremely complex and presents the activation of alternative pathways for energy generation [186]. This results in cell growth alterations, especially a decrease in growth rate. Although parameters such as maintenance of recombinant plasmid and plasmid copy number are primarily related with compatibility of the host and vector system, these are also greatly affected by culture conditions and media composition [187]. The plasmid induced metabolic burden is reflected by alterations in host cell physiology and metabolism caused by stress responses, which has been studied in literature extensively (Table 6.1). In return, this caused lower plasmid DNA yields due to the deleterious effects of these alterations on cell growth.
Table 6.1: Alterations in *E. coli* to plasmid-imposed metabolic burden at cultural and cellular level.

<table>
<thead>
<tr>
<th>Alterations in <em>E. coli</em> to plasmid-imposed metabolic burden</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. At culture level</td>
<td></td>
</tr>
<tr>
<td>Reduced specific growth rates</td>
<td>[188, 189]</td>
</tr>
<tr>
<td>Reduced biomass yield</td>
<td>[189]</td>
</tr>
<tr>
<td>Reduced plasmid stability</td>
<td>[190]</td>
</tr>
<tr>
<td>2. At cellular level</td>
<td></td>
</tr>
<tr>
<td>Cell filamentation</td>
<td>[191, 192]</td>
</tr>
<tr>
<td>Cell viability</td>
<td>[193]</td>
</tr>
<tr>
<td>Cell cycle alterations</td>
<td>[193, 194]</td>
</tr>
<tr>
<td>Activation of stress responses</td>
<td>[195]</td>
</tr>
</tbody>
</table>

### 6.1.2.1 Plasmid Stability

During large scale production of recombinant DNA products, plasmid instability can directly decrease the efficiency and productivity of the fermentation process. Consequently, commercial exploitation of recombinant engineered strains is highly dependent on plasmid stability. To achieve improvement in plasmid stability, it is important to understand the factors affecting plasmid stability and the mechanisms. There are many factors that affect plasmid stability, such as plasmid load, replication patterns, copy numbers of plasmid, type of insert, copy number variance, substrate type, medium composition, host inheritable background, culture conditions growth rate, expression level, growth temperature, and encoded protein toxicity [196-199]. In general, all of these factors can be classified as two aspects, namely, cell internal factors and culture conditions. The former is very important during the construction of recombinant strains. However, for the scaling up and further industrialization of the fermentation process, culture conditions will become the more important aspect.

### 6.1.2.2 Cell Growth

Several research studies have shown that plasmid-bearing cells showed lower specific growth rates than plasmid-free cells, resulting in lower biomass yields at the end of culturing [188, 189]. The level of decrease in specific growth rate is related with plasmid type and copy number [153, 200]. This growth retardation in
plasmid-bearing cells can be attributed to the competition between the two main biological activities that take place in recombinant cells, which are cell propagation and plasmid replication [201].

6.1.2.3 Carbon source

The presence of a carbon source in the medium can repress expression of certain genes and operons. In the presence of glucose, the level of cAMP, which is a component directly stimulates transcription of RNA polymerase, is low. Consequently transcription from the lac promoter is also low. In order to avoid increased basal expression in DE3 lysogens in stationary phase, LB medium can be supplemented up to 10 g/L glucose. For this reason glycerol was used as a main carbon source in this study and the media supplemented 4 g/L of glucose to prevent basal expression that may cause extra burden to the cells in stationary phase.

6.1.3. T7 RNA Polymerase System

The T7 RNA Polymerase System is the most popular approach for producing recombinant proteins in E. coli. In this system, target gene is cloned downstream of the T7 promoter which takes place on an expression vector. After that, the expression vector containing the target gene is introduced into a T7 expression host. T7 expression hosts such as DE3 strains have a chromosomal copy of the phage T7 RNA polymerase gene. When an inducer such as IPTG or rhamnose is added to the culture, this initiates the expression of T7 RNA polymerase and transcription of the target gene, which is followed by translation of the desired protein by endogenous protein translation machinery. A schematic representation of can be seen in Figure 6.1.
6.1.4. Bacterial Growth Phases

The phases of bacterial growth are described in four sections. The length of the lag, exponential and stationary phases of the growth curves vary considerably among the different strains.

**Lag phase:** Lag phase is the adaption period to the new environment after the introduction of a microorganism into a culture medium. In lag phase, cellular metabolism is accelerated, cells are increasing in size, but the bacteria are not able to replicate and therefore no increase in cell mass. The length of the lag phase depends primarily on growth conditions.

**Exponential (log) Phase:** In exponential phase, the cells divide at a constant rate depending upon the composition of the growth medium and the conditions of incubation. The rate of exponential growth of a bacterial culture is expressed as generation time, and also the doubling time of the bacterial population.

**Stationary Phase:** The log phase of bacterial growth is followed by the stationary phase, since the population growth is limited by some factors such as exhaustion of...
available nutrients, accumulation of inhibitory metabolites, exhaustion of space. In stationary phase the size of a population of bacteria remains constant, even though some cells continue to divide and others begin to die.

**Death Phase:** The stationary phase is followed by the death phase, in which the death of cells in the population exceeds the formation of new cells.

![Figure 6.2: Bacterial Growth Phases](image)

### 6.1.5. Aims and Objectives

The aim of the work reported in this chapter is to investigate bacterial cellulose production by genetically modified strain of *E. coli* BL21 (DE3) carrying two expression plasmids: pBCS and pCMP. The objectives of the work reported in this chapter are:

- To monitor the effect of recombinant gene expressions on cell growth and plasmid stability under initial culture conditions (initiated by 0.2 mM IPTG addition, at 37 °C).

- Optimisation of bioprocess conditions (temperature, IPTG concentration, induction OD) for maximum cell growth and plasmid stability.

- To monitor the expression of target genes in the presence of various IPTG concentrations and at various culture temperatures.
• To verify the expression of the target proteins.

• To screen the BC production ability of the genetically modified stains.

The presented chapter is composed of 5 parts: introduction, experimental methodology, results, discussion and conclusion.

6.2. Methodology

The strains and the recombinant plasmids used in this chapter are shown in the following Table 6.2. Other materials used in this chapter characterized in Chapter 4.1.

The experiments reported in this chapter were conducted in effort to examine the effect of temperature and IPTG concentration on cell growth and plasmid stability. For this purpose, the cells were cultured at 37°C and 180 rpm, in LB medium (described in Chapter 4.1.3) supplemented by streptomycin (50 mg/ml), carbenicillin (50 mg/ml) and with 4 g/L glucose to prevent basal expression before IPTG induction. After induction, glycerol was supplemented to media as the carbon source. The experiments were performed in flasks unless otherwise noted. The media was supplemented by various concentrations of IPTG (isopropyl β-D-1-thiogalactopyranoside) for induction of the expression of the target bcs genes. Following IPTG induction, culture temperature shifted to 30°C or 22°C, if necessary. After 24 hour culturing, cells were collected for further analysis.

Samples were collected every 2 hours from media to monitor cell growth and plasmid stability as described in Chapter 4.2.5 and Chapter 4.2.6, respectively. The proportions of the cells maintain the target plasmid were determined by plasmid stability test and calculated in percentages.

Periplasmic and insoluble proteins were extracted to verify the expression of target recombinant proteins after 18 hours after induction by procedure explained in Chapter 4.2.9. Protein concentrations of the protein extracts were determined
using the method described by Bradford (1976) and the samples were analysed by SDS-PAGE. The procedures can be found in Chapter 4.2.9.5 and 4.2.9.6. Protein bands of the samples were compared against the marker to verify the presence of target proteins.

RNA extraction was performed to investigation of the level recombinant gene expression. RNAProtect Bacteria Reagent was used to stabilize RNA is before bacterial cells are lysed. This reagent prevents both degradation of RNA transcripts and induction of genes to ensure reliable gene expression analysis. After cell lysis, the RNeasy Mini Kit or was used to purify total RNA. Details of the procedures are presented in Chapter 4.2.11. Following that, Quantitech Reverse Transcription and QuantiTect SYBR Green PCR Kit were used for reverse transcription and Real Time q-PCR analysis, respectively. Details of the procedures can be seen in Chapter 4.2.15 and Chapter 4.2.16.

Congo red binding assay was used to detect the ability of genetically modified strain to use cellulose for both the cells grown on agar plates and liquid cultures. E. coli BL21 (DE3) (pBCS + pCMP) cells were grown until the OD$_{600}$ reached 1.0 and induced with 0.2 mM IPTG to initiate the gene expression for the production of BC. After that, the cells were diluted and plated as explained in Chapter 4.2.6 on LB agar plate containing 25 mg/L Congo red, 50 mg/ml streptomycin, 50 mg/ml carbenicillin. As control, the E. coli BL21 (DE3) cells grown without IPTG induction also plated on LB agar plate. All plates were incubated at 37 °C for 24 hr and screened qualitatively to detect the BC production ability. Also, E. coli BL21 (DE3) (pBCS + pCMP) induced by 0.2 mM IPTG was grown 18 hours after induction. After that the culture supplied with 25 mg/L Congo red and incubated in static conditions in room temperature and the cells were harvested by centrifugation at 4000 rpm for 10 minutes for qualitative analysis.
6.3. Results

6.3.1. Expression Systems

In this chapter, pBCS carrying bcsABCD operon and pCMP carrying the upstream operon (cmcax and ccpAx) were used for the investigation of cellulose biosynthesis. The experiments were performed by *E. coli* BL21 (DE3) host strain bearing these plasmids, which both work under the control of the strong T7lac promoter. The expression of the recombinant proteins was easily and efficiently achieved by inducing the promoter with various concentrations of IPTG.

<table>
<thead>
<tr>
<th><em>E. coli</em> Strain</th>
<th>Feature</th>
<th>Recombinant proteins synthesized by IPTG induction</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> BL21 (DE3)</td>
<td>Native Strain</td>
<td>-</td>
</tr>
<tr>
<td><em>E. coli</em> BL21 (DE3) (pBCS + pCMP)</td>
<td>Genetically Modified Strain carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CCPAxA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Feature</th>
<th>Expressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBCS</td>
<td>pETDuet-1 with inserted bcsABCD operon encoding cellulose synthase</td>
<td><em>bcs</em>, A, bcsB, bcsC, bcsD,</td>
</tr>
<tr>
<td>pCMP</td>
<td>pCDFDuet-1 with inserted cmcax and ccpAx encoding endo-β-1,4-glucanase and cellulose complementing protein, respectively.</td>
<td>cmcax, ccpAx</td>
</tr>
</tbody>
</table>

In effort to investigate the simultaneous expression of the BC biosynthesis related proteins in *E. coli* BL21 (DE3) (pBCS + pCMP), various IPTG concentrations (0.025, 0.05, 0.1, and 0.2 mM), various temperatures (22, 30, and 37 °C) and
induction OD (1.0, 1.5, and 2.0) were investigated. *E. coli* BL21 (DE3), native strain, was used as control. Expression systems used in the work reported in this chapter and their features are summarized in Table 6.2.

### 6.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of *E. coli* BL21 (DE3)

The expression of recombinant proteins usually imparts a metabolic burden on the microorganism, which causes a considerable decrease on growth rate. Another problem for the expression of recombinant proteins is plasmid instability, which can result in a significant loss in productivity. In effort to investigate those possibilities, growth profiles and plasmid stability profiles of *E. coli* BL21 (DE3) (pBCS + pCMP) cells were investigated. Control groups used in these experiments are summarized in Table 6.3. Also, experimental conditions used in the work reported in this chapter are presented in Table 6.4.

The fermentations were conducted in 500 ml flasks using 100 ml of LB culture to maintain 1:5 LB medium ratio in flasks unless otherwise noted. The flask experiments were performed in triplicate and the bars represent the range of error. Induction time points are indicated by arrows. Each fermentation was repeated three independent times and each analysis was replicated also three times, unless otherwise noted. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to $P < 0.05$. Glycerol was used as a main carbon source and the media supplemented 4 g/L of glucose to prevent the risk of basal expression that may cause extra burden to the cells in stationary phase.

<table>
<thead>
<tr>
<th>Table 6.3: Control and experimental groups used in Chapter 6.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli</strong> BL21 (DE3)</td>
</tr>
<tr>
<td><strong>E. coli</strong> BL21 (DE3) (pBCS + pCMP)</td>
</tr>
<tr>
<td><strong>E. coli</strong> BL21 (DE3) (pBCS + pCMP) + IPTG</td>
</tr>
</tbody>
</table>
The cells were cultured at 37°C, which is the optimum temperature for *E. coli* to grow fast and reach the maximum volumetric productivity. Although 1 mM IPTG is recommended for full induction with vectors having the T7/lac promoter, considering the complexity of the expression system design in this study, lower IPTG concentrations was investigated in the first step. For this reason, 0.2 mM IPTG concentration was chosen as the initial concentration for determination of the effect of inducer concentration on *E. coli* BL21 (DE3) (pBSC + pCMP) cells. When the OD reached 1.0, *E. coli* BL21 (DE3) (pBSC + pCMP) cells were induced by 0.2 mM IPTG. *E. coli* BL21 (DE3) cells and un-induced *E. coli* BL21 (DE3) (pBSC + pCMP) cells were used as controls.

The growth profiles of the cells cultured at 37°C are shown in Figure 6.3. Before IPTG induction, optical densities of genetically modified strains exhibited extremely similar growth profile and reached 0.5 at 600 nm in 5 hours, while the native strain reached 0.5 at 600 nm in less than 2 hours. The reason behind the

Table 6.4: Experimental conditions used in Chapter 6.

<table>
<thead>
<tr>
<th>IPTG (mM)</th>
<th>Temperature (°C)</th>
<th>Fermentation Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>37°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>30°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>22°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.025</td>
<td>22°C</td>
<td>Bioreactor</td>
</tr>
</tbody>
</table>

The growth profiles of the cells cultured at 37°C are shown in Figure 6.3. Before IPTG induction, optical densities of genetically modified strains exhibited extremely similar growth profile and reached 0.5 at 600 nm in 5 hours, while the native strain reached 0.5 at 600 nm in less than 2 hours. The reason behind the
slow growth rate of genetically modified strains is that the cells experience additional burden caused by plasmid maintenance and replication [202].

![Growth profiles of BL21 (DE3) cells at 37°C.](image)

**Figure 6.3:** Growth profiles of BL21 (DE3) cells at 37°C.

After the induction with IPTG 0.2 mM, *E. coli* BL21 (DE3) (pBCS + pCMP) cells (induced cells) exhibited a slower growth profile compared to *E. coli* BL21 (DE3) (pBCS + pCMP) cells that are not induced with IPTG (un-induced cells). This can prove the initiation of the gene expression in the induced cells. However, the growth profile of the genetically modified strains induced with IPTG did not exhibit significant difference as compared with un-induced genetically modified strains at P > 0.05. On the other hand, the growth profile of the genetically modified strains (both induced and un-induced) showed significant difference as compared with native strains at P < 0.05. The genetically modified strains grow slower because of the dedications of all their resources to the replication of the plasmids instead of cell maintenance. In the end of 11 hours fermentation, native cells and the un-
induced cells reached approximately $OD_{600}=2.5$. However cells induced with IPTG stayed in $OD_{600}=1.5$.

Percentage of plasmid-bearing cells was determined to test the stability of the expression system (Figure 6.4). Before IPTG induction, plasmid stability was higher than 80%. 1 hr after IPTG induction, lower than 40% of the induced cells was plasmid-bearing. In the end of fermentation period, the percentage of plasmid-bearing cells among induced group dramatically decreased to 6%. The percentage of plasmid-bearing cells was also low in un-induced control cells with a percentage of 67%. Plasmid stability of the induced cells showed a significant difference as compared with un-induced cells at $P < 0.05$ after IPTG induction.

Figure 6.4: Proportion of the plasmid-bearing BL21 (DE3) cells at 37 °C.
The results showed that although the concentration of IPTG (0.2 mM) used in these experiments was considerably lower (1/5) than recommended concentration of IPTG (1.0 mM), it still caused an extremely high metabolic burden for the cells. Since the induced cells experience additional burden in comparison to un-induced cells caused by plasmid replication and also protein expression, this resulted in slower growth rates and plasmid instability [202].

In effort to reduce the negative effects of induction of IPTG on plasmid stability and cell growth, experiments were conducted in the presence of lower IPTG concentrations (0.025, 0.05, 0.1 mM). The cells were grown at 37°C and when the OD reached 1.0, *E. coli* BL21 (DE3) (pBCS + pCMP) cells induced by IPTG. *E. coli* BL21 (DE3) cells were used as negative control while *E. coli* BL21 (DE3) (pBCS + pCMP) cells without IPTG induction were used as un-induced control.

![Figure 6.5: Growth profiles of BL21 (DE3) cells at 37°C.](image-url)
According to results shown in Figure 6.5, after the induction of \textit{E. coli} BL21 (DE3) (pBCS + pCMP) cells with 0.025, 0.05 and 0.1 mM IPTG, a slower growth profile observed in all induced groups compared to un-induced control group. Also, induced groups exhibited statistically significant difference as compared with native strains at $P < 0.05$. This was expected as a result of metabolic burden caused by the recombinant gene expression. The optical density of the all induced cells reached only approximately 1.5 at 600 nm after 10 hours and the cells entered stationary phase. On the other hand, the negative control cells reached the highest OD at 600 with the value of 2.5, and showed a profile in tendency to reach further OD values.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{image.png}
\caption{Proportion of the plasmid-bearing BL21 (DE3) cells at 37°C.}
\end{figure}
The percentage of plasmid-bearing cells were also monitored and showed in Figure 6.6. 1 hour after induction, a slight difference observed in plasmid stabilities of induced cells in a negative correlation with IPTG concentration. The decrease in the percentage of plasmid-bearing cells after the induction of 0.025 mM and 0.05 mM IPTG was not as sharp as the decrease in the percentage of plasmid-bearing cells induced by 0.2 mM (previous data) and 0.1 mM and IPTG. All induced groups exhibited a significant decrease in plasmid stability in compared with un-induced cells at P < 0.05. None of the induced cells were plasmid-bearing after 10 hours, while 60% of the un-induced cells were bearing plasmid. Any extracellular BC production was not detected in none of the genetically modified groups. This attributed to plasmid instability and protein accumulation in the cells.

Experiments conducted by various IPTG concentrations revealed that lower IPTG concentrations could be used to increase the percentage of plasmid-bearing cells and to promote the production of active proteins. In addition, the use of elevated growth temperatures has been described as one of the reasons for plasmid instability [203]. At high inducing temperatures, the effects of the metabolic burden are eclipsed by the toxicity presented by the production different proteins. For these reasons, the cells were grown at 30°C and induced by various IPTG concentrations to test this hypothesis.
Although the optimal growth temperature for *E. coli* is 37°C, it can tolerate lower temperatures with lower growth rate. After the induction of *E. coli* BL21 (DE3) (pBCS + pCMP) cells with 0.025, 0.05 and 0.1 mM IPTG, a slower growth profile observed in induced groups compared to un-induced groups, exhibiting statistically significant difference as compared with native strains at P < 0.05 (Figure 6.7).

Figure 6.7: Growth profiles of BL21 (DE3) cells at 30°C
The percentage of plasmid-bearing cells were also monitored and showed in Figure 6.8. 1 hour after induction, a slight difference observed in plasmid stabilities of induced groups in a negative correlation with IPTG concentration. However, the profile of percentage of plasmid-bearing cells cultivated at 30°C showed that there was no significant difference in plasmid stabilities between the cells cultured at 30°C and the cells cultured at 37°C. Similarly, all induced groups exhibited a significant decrease in plasmid stability in compared with un-induced cells at P < 0.05. None of the induced cells were plasmid-bearing 5 hours after induction at 30°C, which was also observed in the data obtained at 37°C. Also BC production was not detected in extracellular matrix of the cultures of the genetically modified strains. The cells induced by 0.025 mM IPTG at 30°C and 37°C showed a better
plasmid stability profile compared to the cells induced with 0.05, 0.025 and 0.1 mM IPTG. Also, the proportion of plasmid-bearing cells decreases significantly faster at higher temperature, being plasmids the less maintained between 30°C-40°C [196]. Thus, a combination of low temperature (22°C) and low IPTG concentration (0.025 mM) was selected in effort to improve plasmid stability.

The fermentation was conducted at 37°C until IPTG induction to allow the cells to reach optical density of 1 at 600 nm in 5 hours. After the induction, the temperature shifted to 22°C gradually. The growth and plasmid stability profiles were shown in Figure 6.9 and Figure 6.10, respectively.

![Figure 6.9: Growth profile of BL21 (DE3) cells at 22°C.](image)

The induced cells reached an optical density value of 3 in 10 hours, which was higher than that of 30°C and 37°C. The growth profile of the induced groups did not exhibit significant difference as compared with un-induced or native groups.
at P > 0.05. This could be explained by the decrease in the production of different proteins, which eclipses the effects of the metabolic burden. However, the plasmid-bearing cell proportion among the cells induced with 0.025 mM IPTG were only 5% after 10 hours. In addition, only 55% of the cells that were not induced with IPTG were able to keep the plasmids inside cells.

![Figure 6.10: Proportion of the plasmid-bearing BL21 (DE3) cells at 22°C.](image)

Plasmid stability is also affected by other culture conditions such as dissolved oxygen concentration and pH. A decrease in dissolved oxygen concentration level in growth medium could negatively affect plasmid maintenance [187, 204], since enough dissolved oxygen is required for maintaining plasmid stability. Additionally, a low pH medium causes stress to the bacterial cells, resulting in plasmid loss from the high-density cultures [205]. However, fermentation of the cells in flasks does not allow controlling pH and dissolved oxygen levels. For the purpose of overcoming these problems, a stirred-tank
The bioreactor was used to culture the cells. In order to compare the growth and plasmid stability profiles of the cells cultured in bioreactor with the cell cultured in flasks, a parallel experiment in flask also conducted.

The bioreactor was run at 30 °C until the induction of the cell by 0.025 mM IPTG. Following that, the temperature shifted to 22°C. The pH set at 7.0±0.25. This experiment was replicated two times and average values presented. The DO (dissolved oxygen) probe was calibrated at 0% and set at 100%, obtained using 250 rpm agitation. After calibration, DO remain at approximately 100% until inoculation.

![Graph showing growth profiles](image-url)

**Figure 6.11:** Growth profiles of BL21 (DE3) cells cultured in bioreactor/flask at 22°C.

The growth and plasmid stability profiles are shown in Figure 6.11 and 6.12, respectively. The cells grown in bioreactor reached an optical density of 5.9 in 24 hours, while the optical density was only 2.5 for the cells grown in flask. pH control
was maintained at optimum values (approximately pH 7.0), which resulted in a positive effect on cell growth. However, after induction, a sharp decrease was detected in DO levels, as a result of the increase in the cell density. DO levels could not be maintained at the set-point and fell to very low levels (<2%) in the end of the fermentation. Therefore, the improvement in cell growth was not sufficient. In addition, it was observed that the cells cultured in bioreactor entered a stationary phase after induction, and following that, the cells resumed growth. The loss of the plasmids possibly allowed cells to grow more after induction without the any metabolic burden caused by recombinant gene expression.

Figure 6.12: Proportion of the plasmid-bearing BL21 (DE3) cells cultured in bioreactor/flask at 22°C.
On the other hand, the decrease in the proportion of plasmid-bearing cells cultured in bioreactor and flask showed a similar trend and none of the cells were plasmid-bearing after 9-10 hours without any significant difference. This could be caused by poor DO control, which has a significant effect on plasmid bearing ability of the cells. Also, another reason behind the plasmid instability problem could be related with strain properties. Although expression of a large amount of proteins has been achieved in very high levels in BL21 (DE3), over-production was not successful because of the toxicity of the target proteins, which can also cause cell death. In fact, recombinant proteins that expressed in BL21 (DE3) did not cause cell death, but resulted in plasmid loss. However the expression of six recombinant proteins simultaneously under the control of strong T7lac promoter may require superior strains, which can maintain the plasmids inside the cells.

Induction OD and induction length have also influence on the production of recombinant proteins. The cells should be induced in exponential phase, after the adaptation of the cells to the culture conditions. With the aim of testing the effect of induction OD on cell growth and plasmid stability, the cells were induced by 0.2 mM IPTG, at OD=1.0, 1.5 and 2.0. The cells were grown at 37 °C until the induction, following that, temperature shifted to 30 °C.
Cell growth and plasmid stability profiles were shown in Figure 6.13 and Figure 6.14, respectively. The results showed that the final optical densities were slightly improved when the cells were induced at higher OD. The cells induced at OD= 2.0 reached the highest OD value (2.2), while the cells induced at OD=1.0 and OD=1.5 reached 1.8 and 2.0, respectively. Nevertheless, any significant difference in final OD values was not observed.
Figure 6.14: Proportion of the plasmid-bearing BL21 (DE3) cells induced at OD=1.0, 1.5 and 2.0.

The proportion of plasmid-bearing cells of the groups induced at OD$_{600}$= 1.5 and OD$_{600}$= 2.0 showed a significant improvement for the first 3 hours of induction with a very similar trend. However, only 22% of these groups were plasmid-bearing after 10 hours. This was not satisfactory for a stable and continuous system to accomplish a high-throughput culture. These results also supported that the reason behind the plasmid instability problem could be related with strain properties more than culture parameters [122].
6.3.3. Gene Expression Analysis of BC Synthase Related Genes

In order to confirm expression of the recombinant genes (cmcax, ccpAx, bcsA, bcsB, bcsC, and bcsD) and to determine the effect of IPTG concentration and temperature on the amount of recombinant gene expressions, q-PCR analysis was performed. q-PCR provides a relative value expressed as a ratio of the amount of initial target gene between before IPTG induction (control) and after IPTG induction normalized relative to an internal standard (housekeeping gene). Housekeeping gene used in this analysis was serC, a chromosomal E. coli gene that is usually constitutively expressed [106]. RNA quantification was performed to balance the amount of RNA for each sample before the production of cDNA by reverse transcription. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to P < 0.05.

The cells were cultured at 30°C and harvested for RNA extraction 3, 6 and 18 hours after 0.2 mM IPTG induction. The relative gene expressions were showed in Figure 6.15. The results confirmed that gene expression levels were highly sufficient and significant with P < 0.05, since all recombinant genes were expressed more than $10^3$ fold in 3 hours compared to the amount of gene expression before IPTG induction. The data showed that amount of the gene expression remained approximately at the same levels after 6 hours of induction and then decreased after 18 hours of induction (P < 0.05). Overall, it was observed that expression levels of the genes in each separate time point were in the same range, there was no significant difference obtained (P > 0.05). This indicates that the gene expressions were regulated in balance. The amount of recombinant gene expressions was slightly higher than published data on recombinant gene expression in bacteria [206, 207].
In effort to examine the amount of recombinant gene expression at early periods of induction, the cells were cultured at 30°C and harvested for RNA extraction 15, 30 and 60 minutes after 0.2 mM IPTG induction. The data indicated that all of the target genes expressed significantly 60 minutes after IPTG induction ($P < 0.05$).

High level of gene expressions indicates the high replication levels of plasmids. Since the amount of gene expressions reached more than $10^2$ fold 30 minutes after IPTG induction, this was expected to cause an extreme metabolic burden for the cells, which explains the plasmid instability observed in previous data. In addition, a high level of protein expression may cause accumulation of proteins inside the cells. Since the expression system requires active proteins in correct folding and location, protein accumulation could negatively affect the production of BC.
After it was concluded that 0.2 mM IPTG concentration caused high metabolic burden for the genetically modified cells, the cells were induced by 0.025 mM IPTG at 30°C and harvested for RNA extraction 15, 30 and 60 minutes after induction. According to data shown in Figure 6.17, the amount of the gene expressions induced by 0.025 mM IPTG were lower than the amount of the gene expressions induced by 0.2 mM IPTG, which is more desirable for the production of soluble proteins. Also, the data showed statistically significant increase at P < 0.05. Similar to the previous data, no significant difference observed among the expression levels of the genes at each time point. In order to compare the effect of temperature decrease on gene expression, the cells were cultured at 22°C and induced by 0.025 mM IPTG. This data showed a significant decrease in gene expressions compared to the cells cultured at 30°C (Figure 6.18). There was no gene expression obtained 3 hours after induction (data not shown), while the gene expressions were detected in the end of fermentation (P < 0.05).
**Figure 6.17:** Relative gene expressions of BL21 (DE3) (pBCS + pCMP), 15, 30, 60 min after induction by 0.025 mM IPTG at 30°C.

**Figure 6.18:** Relative gene expressions of BL21 (DE3) (pBCS + pCMP), 18 hr after induction by 0.025 mM IPTG at 22°C.
6.3.4. Congo Binding Analysis

Congo red binding assay was performed to detect the ability of genetically modified BL21 (DE3) strain to produce cellulose. All plates containing 25 mg/L Congo red were incubated at 37 °C for 24 hours and screened qualitatively based on colour change. *E. coli* BL21 (DE3) (pBCS + pCMP) cells were induced with 0.2 mM IPTG and *E. coli* BL21 (DE3) used as negative control.

![Figure 6.19: Congo binding analysis of BL21 (DE3) (1) and BL21 (DE3) (pBCS + pCMP) (2) on agar plates.](image)

Colonies containing the plasmid vectors, which can be seen in Figure 6.19 (2), for cellulose production turned red within 24h, and presented a rougher composition and brighter red color. In contrast to this, colonies not containing the plasmid vectors (negative controls), which can be seen in Figure 6.19 (1), presented no color changes and maintained the characteristics of standard *E. coli* colonies (smoother composition and standard yellow colour). As a result of metabolic burden caused by induction, colonies containing plasmid vectors were smaller than the negative controls.
In effort to confirm this qualitative assay, a second qualitative assay was conducted, which was also based on Congo Red binding. *E. coli* BL21 (DE3) (pBCS + pCMP) induced by 0.2 mM IPTG was grown 18 hours after induction. Following that, the culture supplied with 25 mg/L Congo red and incubated in static conditions in room temperature and the cells were harvested by centrifugation at 4000 rpm for 10 minutes for qualitative analysis. Cellulose producers were expected to grow cellulose fibres on their extracellular surface. Consequently, this should present a colour change in cellulose producing cells in the presence of Congo Red, unlike non-cellulose-producing cells. After harvesting the cells, a red pellet was observed that corresponded to the cellulose producing colonies (Figure 6.20 (1)). In contrast to this, colour change was not observed in negative control cells (Figure 6.20 (2)). These results indicated that the cellulose production operon has been successfully refactored and transferred into *E. coli*.

### 6.3.5. SDS PAGE Analysis

Since BC is formed between the outer and cytoplasm membranes of the cell [37], the recombinant proteins are located either in periplasm or embedded in between membranes. Cmcax is secreted into growth medium by bacteria [47], while CcpAX co-localizes with BcsD [208], which is a periplasmic protein [209, 210]. BcsA is an integral membrane protein that contains the catalytic domain and forms a trans-membrane (TM) pore across the inner membrane. BcsB is a periplasmic
protein which guides the polymer across the periplasm toward the outer membrane [210]. BcsC is a pore-forming outer membrane protein.

*E. coli* BL21 (DE3) (pBCS + pCMP) induced by 0.2 mM IPTG was grown 3 hours after induction at 37°C. The proteins extracted from induced cells were compared with proteins extracted from native strain (*E. coli* BL21 (DE3)) and un-induced cells (*E. coli* BL21 (DE3) (pBCS + pCMP)). In order to detect the proteins by SDS-PAGE analysis, periplasmic proteins and insoluble fractions (membrane proteins) were extracted as described in Chapter 4.2.9. Concentration of the total proteins in extracts was determined by Bradford Assay as described in Chapter 4.2.9.5. SDS-PAGE analysis was performed as explained in Chapter 4.2.9.6. The gels are presented in Figure 6.19.

The protein sizes of CcpAx, Cmcax, BcsA, BcsB, BcsC and BcsD are 39.0, 37.2, 80.0, 89.6, 144.8 and 18.5 kDa, respectively. In Figure 6.21 (A), Line 1 shows the molecular weight marker, while Line 2, 3, 4 shows periplasmic extracts of native cells, un-induced cells and induced cells, respectively. In Figure 6.21 (B), Line 1 shows the molecular weight marker, while Line 2, 3, 4 shows insoluble extracts of native cells, un-induced cells and induced cells, respectively. A band corresponding size of BcsC was detected in the insoluble fraction since it is a membrane associated protein, while the other proteins were detected on periplasmic fraction. Detection of an intense protein band corresponding to size of BcsC in insoluble fraction could be also due to the inclusion body formation. CcpAX and Cmcax were detected on the gel slightly upper than the region that correspond the sizes of these proteins. This could be caused by existence of other sequences on the cloning side either upstream of the protein after translation starts or downstream of the protein before translation ends. Translation of Cmcax starts with a sequence including His-tag with the size of approximately 5 kDa located its upstream. Similarly, translation of CcpAX starts with a 2.5 kDa sequence. These size differences could cause a ‘shifting’ in the protein bands, which is common for especially small size proteins, as they have higher mobility on the gel [211]. For this reason, both proteins could be expected to visualize higher than 42 kDa.
Figure 6.21: SDS-PAGE analysis of periplasmic (A) and insoluble (B) proteins extracted from *E. coli* BL21 (DE3) (pBCS + pCMP).

Total intracellular proteins and cytoplasmic fractions were also extracted and analysed by SDS analysis. However these fractions contain high concentrations of entire proteins taking part in metabolism of a cell. For this reason, many overlapped proteins and extremely dense bands were observed in a single lane in the gels. Consequently, any significant difference was not observed between protein extracts obtained from induced and un-induced cell groups, since the protein bands and overlapped each other (data not presented).

### 6.4. Discussion

Economic feasibility of bacterial cellulose is primarily related with its productivity. Relatively high cost of the production of cellulose limits its application to high value-added products such as artificial skin, high fidelity acoustic speakers, and high quality paper. *A. xylinum* has a long doubling time (4 hour in agitated, 8 hours in static culture) compared to most other bacteria, such as *E. coli* (20 minutes). Since its growth rate is significantly faster than that of *A. xylinum*, *E. coli* was utilised to investigate BC production. The availability of a large number of...
cloning vectors and various mutant strains ensure that *E. coli* strains are favoured as host microorganisms for recombinant protein expression [186].

In this Chapter, recombinant plasmids pBCS and pCMP, which are compatible with strain BL21 (DE3), transformed into bacteria in effort to create a cellulose producing *E. coli* strain. The strain BL21 (DE3) is the most widely used strain for protein expression, because it is deficient in the production of proteases Lon and OmpT, which decreases the breakdown of recombinant protein and produces lower amount of acetate compared with other *E. coli* strains results in higher yields[117, 118, 185]. However, expression of six different genes simultaneously is very likely to result in metabolic burden for the genetically modified cells. A number of biological restraints initiated by plasmid maintenance and replication as well as culture conditions are responsible for limiting final biomass and product yields. Culture temperature is one the most important culture parameters for plasmid stability and protein solubility. Protein production at lower temperatures has a profound impact on protein quality. It has been shown that conformational quality and functionality of highly soluble recombinant proteins increase when the temperature of the culture is reduced [192, 212]. Target protein yields are increased several-fold using IPTG induction [213]. Expression of proteins can be regulated by the optimisation of IPTG concentration, as it can dramatically influence expression. For the expression of some proteins it is important to slowly induce transcription of the expression plasmid (with lower IPTG concentrations) since lower levels of expression can increase the solubility and activity of some target proteins [214]. Therefore, various IPTG concentrations (0.025, 0.05, 0.1, and 0.2), various temperatures (22, 30, and 37 °C) and induction OD (1.0, 1.5, and 2.0) were investigated.

The optimal growth temperature for *E. coli*, 37°C, was selected for fermentation and various IPTG concentrations (0.025, 0.05, 0.1 mM) was used to induce the gene expression in genetically modified cells. However none of the genetically modified cells was plasmid-bearing after 10 hours and cellulose production was not detected. Schein and coworkers reported that growth at 37°C
caused some proteins to accumulate as inclusion bodies, while incubation at 30°C lead to soluble and active protein. For this reason, *E. coli* BL21 (DE3) (pBCS + pCMP) cells were grown at 30°C and induced with 0.025, 0.05 and 0.1 mM IPTG. None of the induced cells were plasmid-bearing 5 hours after induction at 30°C. A slower growth profile observed in all induced cells compared to un-induced control and negative control group at both 30°C and 37°C. It is known that overexpression of membrane proteins can lead to non-optimal cell growth [120]. The cells induced by 0.025 mM IPTG at 30°C and 37°C showed an improved plasmid stability profile compared to the cells induced with 0.05, 0.025 and 0.1 mM IPTG. Prolonged (e.g., overnight) induction at low temperatures (15-20°C) may prove optimal for the yield of soluble protein [207]. For this reason, *E. coli* BL21 (DE3) (pBCS + pCMP) cells were grown at 30°C and induced with 0.025. The induced cells reached an optical density value of 3 in 10 hours, which was higher than that of 30 °C and 37 °C. This might be explained by the decrease in the production of different proteins, which reduces the effects of the metabolic burden. However, the plasmid-bearing cell proportion among the cells induced with 0.025 mM IPTG were only 5% after 10 hours and BC production was not observed. On the other hand, the decrease in the proportion of plasmid-bearing cells cultured in bioreactor and flask showed a similar trend and none of the cells were plasmid-bearing after 9-10 hours. O₂ supply in bioreactor was not efficient to maintain the DO level at the set point. Therefore, neither growth profiles nor plasmid stabilities could be improved. Also, growth and plasmid stability profiles obtained by the induction of the genetically modified cells at different optical density values were not satisfactory.

Gene expression levels were also investigated under different conditions, which showed there was no significant difference between the expression levels among the genes in the same time point (P > 0.05). This strongly supports that the expression levels of the genes are in stoichiometric balance. In effort to examine the amount of recombinant gene expression at early periods of induction, the cells were cultured at 30°C and harvested for RNA extraction 15, 30 and 60 minutes after 0.2 mM IPTG induction. The data indicated that all of the target genes expressed significantly even 15 minutes after IPTG induction. For effective expression of
membrane proteins, transcription of the expression plasmids should be slowly induced with lower IPTG concentrations, to increase the solubility and activity of some target proteins [101, 214]. The high gene expression levels at 30˚C indicated that BC production was not obtained as a result of protein accumulation. Detection of a highly dense BcsC in the SDS gel analysis in insoluble fraction supports that theory. Although cellulose production ability of the genetically modified stain was tested and confirmed by Congo binding analysis, BC production was not detected in cultures. This also supports the likeliness of the accumulation of inactive BcsC, which plays a critical role in the regulation of BC export. Lowering the temperature and IPTG concentration was not able to solve the dramatic plasmid loss of the expression system. This showed that although BL21 (DE3) strain was compatible with pBCS and pCMP, BL21 (DE3) encountered plasmid loss [122]. The resulting strong and high-level protein expression can overburden the metabolic capacities of the production host and lead to growth inhibition, which was obtained in our system (highest OD$_{600}$=3.0) [76, 215]. Therefore, the actual production phase during a process is very limited [122].

Strong host/vector interactions also contribute to the stress level of the host. Effect of recombinant protein production onto the host metabolism has been draw attention of scientist, as the host-vector interactions during severe conditions are extremely complex and unpredictable. Although expression of a large amount of proteins has been achieved in very high levels in BL21 (DE3), over-production was not successful because of the toxicity of the target proteins. Considering that the fact that most of the membrane proteins are toxic, overproduction of the target proteins could cause cell death and plasmid instability. Despite the fact that stain BL21 (DE3) is a workhorse for overexpression of recombinant proteins, it was demonstrated that this stain is not favourable for a balanced co-expression of active proteins.

Marisch and colleagues compared the plasmid retaining and protein expression capacities of strain BL21 (DE3) and strain HMS174 (DE3). The host strain BL21 (DE3) produced the higher amounts of specific protein, however expression
system HMS174 (DE3) exhibited a highly stable system by retaining the expression vector over the entire process time (100% after 14 hours), but the cells grew slowly after induction. In contrast, BL21 (DE3) encountered plasmid loss but maintained growth[122].

### 6.5. Conclusion

The expression of toxic proteins in *E. coli* strains causes plasmid instability and plasmid-free cells overgrow those retaining it. Although recombinant proteins that expressed in BL21 (DE3) in this study did not cause cell death, it was resulted in an inhibition in growth and a dramatic plasmid loss. Expression of six recombinant proteins simultaneously under the control of strong T7lac promoter may require superior strains, which can maintain the plasmids inside the cells. As the expression system hosted by HMS174 (DE3) cells exhibited a high plasmid in literature, this strain was selected to continue for the further investigation of BC production in *E. coli*. 
7. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED E. COLI STRAIN 2: HMS174 (DE3)
7.1. Introduction

7.1.1. *E. coli* HMS174 (DE3) as a Host Microorganism

As mentioned in Chapter 6.1.2, the choice of the host cell as for the production of the target proteins defines the technology, molecular tools, equipment, or reagents required for the project. An investigation in the literature for a suitable *E. coli* strain for the selection as a host would produce a large amount of potential candidates. Most of them are special strains that are used in particular situations.

The most extensively used laboratory strains of *E. coli* have been derived from K-12 and B derivatives. Cultivation characteristics exhibited great differences among the K-12 derivatives and B derivatives, as the B derivatives accomplished higher growth rates, higher biomass yields, and lower acetate accumulation. On the other hand, HMS174 (DE3), which is a strain derived from K-12, showed a great potential with a very high plasmid stability.

Substrain HMS174 is an engineered successor of the *E. coli* K-12 wild-type strain [121]. Similar to its ancestor, it is an important organism in biotechnological investigations and is used in fermentation processes for heterologous protein production. HMS174 was originally used for the improvement of the structural and genetic map of the T7 phage in the Studier lab at Brookhaven [184]. After that, it became a popular strain for the expression of recombinant proteins after lysogenization with the DE3 prophage.

7.1.2. Effects of Plasmid Presence on *E. coli* HMS174 (DE3)

Plasmid DNA maintenance and replication commonly induces a “metabolic burden” in *E. coli*. In return, this caused lower plasmid DNA yields and decrease in cell growth. It has been proved that plasmid-bearing cells showed lower specific growth rates than plasmid-free cells, resulting in lower biomass yields at the end of culturing [188, 189].
**E. coli** HMS174 provides a recA1 mutation in a K-12 background. This mutation has a positive influence on plasmid stability [122]. Due to this mutation, certain target genes whose products may cause the loss of the DE3 prophage become stabilized. As a consequence, it shows very high plasmid stability.

Marisch *et al.* studied the comparison of BL21 and HMS174 in order to identify individual benefits of each host cell line, as well as weaknesses of the operated process conditions. Both hosts were unable to maintain their growth rate after induction because the recombinant gene expression rate surpassed the physiologically tolerable limits. *E. coli* BL21 later recovered their cell growth, likely due to plasmid loss. In contrast, *E. coli* HMS174 did not resume cell growth due to its apparent higher system stability. In terms of productivity, produced larger both strains produced similar amounts of soluble proteins, however HMS174 would be the candidate of choice concerning plasmid stability and system integrity [122].

In this chapter, as a consequence of plasmid instability and inclusion body formation observed in BL21 (DE3), HMS174 (DE3) was selected as a host microorganism, which exhibits great potential with its high plasmid stability.

### 7.1.3. Cellulose production in *E. coli*

Although BC used in industrial mainly produced by bacterial species such as the model organism *A. xylinus*, BC producing bacteria is not restricted to acetic acid bacteria. Zogaj *et al.* demonstrated that cellulose biosynthesis in *E. coli* and in *Salmonella* occurs with the production of thin aggregative fimbriae, the second component of the extracellular matrix of a multicellular morphotype [116]. Curli fimbriae, cellulose and cell aggregations are components in *Salmonella* biofilms as well as pathogenic and commensal *Escherichia coli* strains [216, 217].

Multicellularity offers strategic advantages to bacteria at specific growth phases compared to single cells [218]. Multicellular behaviour in *Salmonella typhimurium* (rdar morphotype) showed that at least two extracellular matrix components are produced, thin aggregative fimbriae (agf) and cellulose. Recently,
adrA (agfD regulated gene) was shown to be involved in the cellulose production [219]

The expression of matrix components of the rdar morphotype can detect by Congo Red binding assay. Expression of thin aggregative fimbriae leads to a brown colony (bdar morphotype), while expression of cellulose exhibits a pink colony (pdr morphotype). The features of multicellular behaviour could be attributed to the specific components of the extracellular matrix. Thin aggregative fimbriae facilitate adhesion to abiotic surfaces (biofilm formation) and fragile cell-cell interactions, whereas expression of the cellulose leads to a reduced biofilm formation and elastic cell-cell interactions.

Expression of cellulose requires the bcsABZC operon encoding structural genes for cellulose biosynthesis in E. coli [116, 220]. Although the cellulose biosynthesis genes bcsA, bcsB, bcsZ and bcsC (formerly yhjO, yhjN, yhjM and yhjL) are constitutively expressed, cellulose synthesis occurs only when adrA is expressed. adrA encodes a putative transmembrane protein regulated by agfD. Homologous regulatory genes of cellulose biosynthesis in E. coli and A. xylinum were shown in Figure 7.1. Since the above mentioned genes involved in cellulose biosynthesis are highly conserved in E. coli, it is expected that a similar regulatory mechanism leading to cellulose biosynthesis in this microorganism [31].
The biosynthesis of curli fimbriae is achieved by two operons: csgDEFG and csgABC. Curli fimbriae are composed of CsgA, which is the structural protein subunit. CsgD, is required for the activation of curli and also cellulose biosynthesis. This genetic network, which allows rdar morphotype expression, has mostly been studied in Salmonella Typhimurium ATCC14028, while regulation of curli expression was also investigated in E. coli K-12 strains [115, 116, 219, 221].

Each of these substances exhibits a particular type of cell-cell and also cell-surface interaction [219]. Thin aggregative fimbriae form rigid, but fragile connections between cells. In contrast, cellulose forms a connection between the cells through elastic, but stable bonds.

Existence of cellulose biosynthesis genes in E. coli was demonstrated by SofiaBurlandDanielsPlunkettBlattner [220]. However their function has remained to be elucidated since cellulose production is abolished in E. coli K-12 laboratory strains. The expression of curli has been studied by the laboratory strain E. coli MC4100, a K-12 derivative, however the strain MC4100 showed a bdar-like morphotype and cellulose production was not detected [221]. On the other hand E. coli ECOR 10, E. coli ECOR 12, E. coli TOB1 showed a rdar-like morphotype and produced cellulose [116].

So far, cellulose production was detected in pathogenic and commensal strains of E. coli in microscopic amounts and also in a combination with other substances. For these reasons, a non-pathogenic and high-throughput production system is required for the industrial use.
7.1.4. Aims and Objectives

The aim of the work reported in this chapter is to investigate bacterial cellulose production by genetically modified strain of *E. coli* HMS174 (DE3) carrying two expression plasmids: pBCS and pCMP. The objectives of the work reported in this chapter are:

- Optimisation of bioprocess conditions (temperature, IPTG concentration) for maximum cell growth and plasmid stability.
- To verify the expression of target genes and proteins in genetically modified strain.
- To screen the BC production ability of the genetically modified stains.
- To characterise the product by SEM and FTIR.

The presented chapter is composed of 5 parts: introduction, experimental methodology, results, discussion and, conclusion.

7.2. Methodology

The strains and recombinant plasmids used in this chapter are shown in the following Table 7.1. Other materials used in this chapter characterized in Chapter 4.1.

The experiments reported in this chapter were conducted in order to investigate the effect of temperature and IPTG concentration on cell growth and plasmid stability. For this purpose, the cells were cultured at 37°C and 180 rpm, in LB medium (described in Chapter 4.1.3) supplemented by streptomycin (50 mg/ml), carbenicillin (50mg/ml) and with 4 g/L glucose to prevent basal expression before IPTG induction. The experiments were performed in flasks unless otherwise noted. The media was supplemented by various concentrations of IPTG when the optical density reached 1.0 at 600 nm for induction of the expression of the target *bcs*
genes. Following IPTG induction, media was supplemented with glycerol as carbon source culture temperature shifted to 30°C or 22°C, if necessary. After 24 hour culturing, cells were collected for further analysis.

Samples were collected every 2 hours from media to monitor cell growth and plasmid stability as described in Chapter 4.2.5 and Chapter 4.2.6, respectively. The proportions of the cells maintain the target plasmid were determined by plasmid stability test and calculated in percentages. The colonies were counted to determine the colony forming units on each Petri dish. Average values from triplicates plates were taken. Proportions of the viable cells maintain the plasmids were calculated in percentages.

Periplasmic and insoluble proteins were extracted to visualize and confirm the expression of target recombinant proteins after 18 hours after induction by procedure explained in Chapter 4.2.9. Protein extracts were analysed by SDS-PAGE, by following the protocol presented in Chapter 4.2.9.6. The protein bands of the samples were compared against the marker to verify the presence of target proteins.

RNA extraction was performed for the confirmation of the recombinant gene expression. RNAProtect Bacteria Reagent was used to stabilize RNA is before bacterial cells are lysed in order to prevent the degradation of RNA transcripts. Following the cell lysis, the RNeasy Mini Kit or was used to purify total RNA. Details of the procedure can be seen in Chapter 4.2.11. Following that, Quantitech Reverse Transcription and QuantiTect SYBR Green PCR Kit were used for reverse transcription and Real Time q-PCR analysis, respectively. Details of the procedures can be seen in Chapter 4.2.15 and Chapter 4.2.16.

Congo red binding assay was performed to detect the ability of genetically modified strain to use cellulose for both the cells grown on agar plates and liquid cultures. The assay is described in Chapter 4.2.19.

FTIR spectroscopy was used to identify the functional groups in cellulose samples produced by bacteria. ATR-IR spectra were recorded using a Spectrum One
FTIR-spectrometer (Perkin Elmer, Massachusetts, USA). The spectra were collected at a resolution of 2 cm\(^{-1}\) in the range of 600 and 4,000 cm\(^{-1}\).

The morphology and the microstructural features of samples were also investigated using a field emission scanning electron microscope (FESEM), LEO 1525 (Zeiss, Germany) operating at 5kV. Prior to analysis, samples were dried until constant weight and placed on a double-sided carbon tape mounted onto an aluminium stub. They were gold coated for 2 min at 20mA using an Emitech K575X Peltier (Ashford, UK) cooled sputter coater.

7.3. Results

7.3.1. Expression Systems

In this Chapter, the experiments were performed by *E. coli* HMS174 (DE3) host strain bearing the expression plasmids: pBCS carrying *bcsABCD* operon and pCMP carrying the upstream operon (*cmca* and *ccpA*). The expression of the recombinant proteins was easily and efficiently achieved by inducing the strong T7lac promoter with IPTG.

In effort to investigate the BC biosynthesis and the effect of simultaneous expression of target proteins on growth profile and plasmid stability of *E. coli* HMS174 (DE3) (pBCS + pCMP), the cells were induced by 0.025 or 0.2 mM IPTG at various temperatures (22, 30, and 37 °C). *E. coli* HMS174 (DE3) cells, native strain, were used as control. Expression systems used in the work reported in this chapter and their features are summarized in Table 7.1.
Table 7.1: *E. coli* strains and plasmid vectors used in Chapter 7.

<table>
<thead>
<tr>
<th><em>E. coli</em> Strain</th>
<th>Feature</th>
<th>Recombinant proteins synthesized by IPTG induction</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> HMS174 (DE3)</td>
<td>Native Strain</td>
<td>-</td>
</tr>
<tr>
<td><em>E. coli</em> HMS174 (DE3) (pBCS + pCMP)</td>
<td>Genetically Modified Strain carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CcpAx</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Feature</th>
<th>Expressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBCS</td>
<td>pETDuet-1 with inserted <em>bcsABCD</em> operon encoding cellulose synthase</td>
<td><em>bcsA, bcsB, bcsC, bcsD</em>,</td>
</tr>
<tr>
<td>pCMP</td>
<td>pCDFDuet-1 with inserted <em>cmcax</em> and <em>cppAx</em> encoding endo-β-1,4-glucanase and cellulose complementing protein, respectively.</td>
<td><em>cmcax, cppAx</em></td>
</tr>
</tbody>
</table>

7.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of *E. coli* HMS174 (DE3)

As mentioned in Chapter 6, the expression of recombinant proteins usually imparts a metabolic burden on the microorganism, which causes a considerable decrease on growth rate. Also, it may cause plasmid instability, which can result in a significant loss in productivity. In effort to investigate those possibilities, growth a plasmid stability profiles of *E. coli* HMS174 (DE3) (pBCS + pCMP) cells were investigated. Control groups used for these experiments are summarized in Table 7.2. Also, experimental conditions used in the work reported in this chapter are summarized in Table 7.3.
Table 7.2: Control and experimental groups used in Chapter 7.

<table>
<thead>
<tr>
<th>E. coli HMS174 (DE3)</th>
<th>Negative Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli HMS174 (DE3) (pBCS + pCMP)</td>
<td>Un-induced Control</td>
</tr>
<tr>
<td>E. coli HMS174 (DE3) (pBCS + pCMP) + IPTG</td>
<td>Induced Group</td>
</tr>
</tbody>
</table>

In Chapter 6, it is reported that when E. coli BL21 (DE3) (pBCS + pCMP) cells were cultured at 30°C and 37°C, they exhibited a dramatic decrease in the percentage of plasmid-bearing cells after IPTG induction. The percentages of plasmid-bearing cells cultured at 30°C and 37°C were both lower than 20%, 3 hours after 0.025 mM IPTG induction, while it was 40% at 22°C. Plasmid stability and growth profile of E. coli HMS174 (DE3) (pBCS + pCMP) cells were also investigated at 22°C, 30°C and 37°C in order to collect comparable data with the data obtained on E. coli BL21 (DE3) (pBCS + pCMP).

Table 7.3: Experimental conditions used in Chapter 7.

<table>
<thead>
<tr>
<th>IPTG (mM)</th>
<th>Temperature (°C)</th>
<th>Fermentation Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>37°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>30°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>22°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.025</td>
<td>22°C</td>
<td>Bioreactor</td>
</tr>
</tbody>
</table>
The fermentations were conducted in 500 ml flasks using 100 ml of LB culture unless otherwise noted. The experiments were performed in triplicate and the bars represent the range of error. Each fermentation was repeated three independent times and each analysis was replicated also three times, unless otherwise noted. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to P < 0.05. The growth profiles were monitored at 600 nm. Induction time points are indicated by arrows. Glycerol was used as a main carbon source and the media supplemented 4 g/L of glucose to prevent basal expression that may cause extra burden to the cells in stationary phase.

Strain HMS174 (DE3) was cultured first at 37˚C, when the OD value reached 1.0; the cells were induced by 0.025 mM or 0.2 mM IPTG (Figure 7.2). E. coli HMS174 (DE3) cells and un-induced E. coli HMS174 (DE3) (pBCS + pCMP) cells were used as control groups.

![Growth profiles of HMS174 (DE3) cells at 37°C.](image)

*Figure 7.2:* Growth profiles of HMS174 (DE3) cells at 37˚C.
Before IPTG induction, optical densities of the cells exhibited a similar growth trend and reached an optical density of approximately 1.0 at 600 nm in 4 hours 15 minutes. After induction, induced cells entered the stationary phase while the native cells showed a rapid growth and reached an optical density of 2.6 at 600 nm. There was no significant difference between the growth profiles of induced groups (P > 0.05). The lowest optical density (OD$_{600} = 1.1$) was observed in the cells induced by 0.2 mM concentration of IPTG.

![Figure 7.3: Proportion of the plasmid-bearing HMS174 (DE3) cells at 37 °C.](image)

Percentage of plasmid-bearing cells was determined to test plasmid stability (Figure 7.3). Before IPTG induction, plasmid stabilities were observed between 97-88%. The decrease in the proportion of plasmid bearing cells of induced group with
0.025 mM IPTG was not as sharp as the decrease observed in induced group with 0.2 mM IPTG, 1 hour after the induction. Nevertheless, after 9 hours, the proportion of plasmid bearing cells among both induced groups approximately 3%. Although the response of the induced HMS174 (DE3) (pBCS + pCMP) cells showed a similar trend with that of BL21 (DE3) (pBCS + pCMP), un-induced HMS174 (DE3)(pBCS + pCMP) cells showed an improved plasmid stability profile (91%) in 10 hours culturing, which was found promising for further investigations. The cells were cultured for further 14 hours at 90 rpm to investigate the BC production; however any extracellular substance was not obtained in the cultures under these conditions.

Figure 7.4: Growth profiles of HMS174 (DE3) cells at 30°C
Fermentation of the cells was performed at 30°C and induced by 0.025 IPTG or 0.2 mM IPTG in effort to improve the growth and plasmid stability. According to data presented in Figure 7.4, the growth of the induced cells was improved compared to the cells cultured at 37°C. The final OD of the induced cells cultured at 37°C was between 1.1-1.3, whereas it is approximately 1.85 for the induced cells cultured at 30°C. No significant difference was observed between the growth profiles of induced groups (P > 0.05).

![Figure 7.5: Proportion of the plasmid-bearing HMS174 (DE3) cells at 30 °C.](image)

A significant improvement was observed in the proportion of plasmid-bearing cells in induced groups cultured at 30°C. The final percentage of plasmid-bearing cells induced by 0.025 mM IPTG and 0.2 mM IPTG were detected as 97% and 82%, respectively. Although BC production was not detected under these
conditions, this data showed the potential of this strain for the further investigations on production of BC.

Fermentation of the induced cells was performed also at 22 °C. The cells were grown at 37 °C until IPTG induction to allow the cells to reach optical density of 1 at 600 nm in 4 hours. After the induction, the temperature shifted to 22 °C gradually. The growth and plasmid stability profiles were shown in Figure 7.3 and Figure 7.5, respectively.

![Growth profiles of HMS174 (DE3) cells at 22°C.](image)

**Figure 7.6**: Growth profiles of HMS174 (DE3) cells at 22°C.

When the cells cultured at 22 °C, extremely similar growth profiles were observed for the all cell types, no significant difference was detected between the growth profiles of native and induced groups (P > 0.05). The induced cells were reached OD<sub>600</sub> = 2.2 after 10 hours. Moreover, the plasmid-bearing cell proportion
among the cells induced with 0.025 mM IPTG remained remarkably as high as that of un-induced groups, which were both above 85% after 10 hours fermentation. This data proved that, at low temperatures, the decrease in the production of different proteins eclipses the effects of the metabolic burden. The cells were cultured for further 14 hours at 90 rpm. For the first time, cellulose production was achieved under these conditions with the improvement of plasmid stability.

![Figure 7.7: Proportion of the plasmid-bearing HMS174 (DE3) cells at 22 °C.](image)

DO concentration and pH are important bioprocess parameters that are effective on plasmid stability and protein expression [222]. The optimal DO concentration for the specific activity of recombinant proteins was dependent upon the strain [223]. Additionally, a low pH medium causes stress to the bacterial cells, resulting in plasmid loss from the high-density cultures [181]. In effort to control pH
and DO in cultures, the cells were cultured in bioreactor. The bioreactor was run at 30 °C until the induction of the cell by 0.025 mM IPTG. Following that, the temperature shifted to 22°C. The pH set at 7.0±0.25. The dissolved oxygen (DO) was calibrated at 0% and set at 100%, obtained using 250 rpm agitation. After calibration, DO remained at approximately 100% until inoculation. In order to compare the growth and plasmid stability profiles of the cells cultured in bioreactor with the cell cultured in flasks, a parallel experiment in flask also performed. This experiment was replicated two times and average values presented.

The growth profiles are shown in Figure 7.8. Contrary to the results obtained by culturing BL21 (DE3) (pBSS + pCMP), HMS174 (DE3) (pBSS + pCMP) cells cultured in bioreactor did not exhibit an improved growth profile compared to the cells cultured in flask. The final optical density of HMS174 (DE3) (pBSS + pCMP) cells

![Figure 7.8: Growth profiles of HMS174 (DE3) cells cultured in bioreactor/flask at 22°C](image-url)
cultured in bioreactor was detected as lower than 2.0 at 600 nm, however, this was approximately 6.0 in BL21 (DE3) (pBCS + pCMP) cells. This could be caused by the poor O₂ supply for the culture, as the level of DO fell rapidly after induction and it was obtained lower than approximately 2% in the end of fermentation. The combination of poor growth kinetics of strain HMS174 and poor O₂ supply resulted in insufficient growth.

On the other hand, the percentage of plasmid bearing cells cultured in bioreactor was slightly higher than that of the cell cultured in flasks, which is possibly as a result of control of oxygen supplement. High plasmid stability (above 80%) was detected even after 10 hours in both cultures. However BC production was not detected in bioreactor.

![Figure 7.9: Proportion of the plasmid-bearing HMS174 (DE3) cells cultured in bioreactor/flask at 22°C.](image)
7.3.3.  Gene Expression Analysis of BC Synthase Related Genes

q-PCR analysis was performed in order to confirm expression of the recombinant genes (*cmcax, ccpAx, bcsA, bcsB, bcsC and bcsD*) cellulose producing HMS174 (DE3) (pBCS + pCMP) cells. The housekeeping gene used in this analysis was *serC*, a chromosomal *E. coli* gene that is usually constitutively expressed [106]. RNA quantification was performed to balance the amount of RNA for each sample before the production of cDNA by reverse transcription.

The cells were cultured at 22°C and harvested for RNA extraction 3 and 18 hours after 0.025 mM IPTG induction. The relative gene expressions were showed in Figure 7.10. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to P < 0.05.

In Chapter 6, it was reported that the gene expressions was detected 18 hours after induction of BL21 (DE3) (pBCS + pCMP) by 0.025 mM IPTG at 22°C. Contradictory to this data, gene expressions in HMS174 (DE3) (pBCS + pCMP) were detected earlier (3 hours after induction), in a low and significant level (between 1.41-2.71 fold), which is more desirable for the production of soluble proteins. After 18 hour induction, the data demonstrated that the relative expression of the genes was lower (1.03-2.01 folds). The low levels of gene expressions could results in elimination of inclusion body accumulation inside cells and allow the production of active proteins.
Congo Binding Analysis

Congo red binding assay was performed to detect the ability of genetically modified HMS174 (DE3) strain to produce cellulose. All plates containing 25 mg/L Congo red were incubated at 37 °C for 24 hours and screened qualitatively to detect the BC production ability. *E. coli* HMS174 (DE3) (pBCS + pCMP) cells were induced with 0.2 mM IPTG and *E. coli* HMS174 (DE3) used as negative control.

Colonies bearing the plasmid vectors for cellulose production turned red within 24 hours, and presented a rougher composition and brighter red colour (Figure 7.11, (2)). On the other hand, plasmid-free (negative controls) cells presented no colour changes and maintained the characteristics of standard *E. coli* colonies (smoother composition and standard yellow colour) (Figure 7.11 (1)). As a result of metabolic burden caused by induction, colonies bearing plasmid vectors were smaller than the negative controls.
In effort to confirm this qualitative assay, a second qualitative assay was performed, which was also based on Congo Red binding. *E. coli* HMS174 (DE3) (pBCS + pCMP) induced by 0.2 mM IPTG was grown 18 hours after induction. Following that, the culture supplied with 25 mg/L Congo red and incubated in static conditions in room temperature and the cells were harvested by centrifugation at 4000 rpm for 10 minutes for qualitative analysis. Cellulose producers was expected to grow cellulose fibres on their extracellular surface. As a result, this should present a colour change in cellulose producing cells in the presence of Congo Red, unlike non-cellulose-producing cells. After harvesting the cells, a red pellet was observed that corresponds to the cellulose producing colonies (Figure 7.12 (1)). In contrast to this, the colour changed slightly in negative control cells, possibly arise from the accumulation of Congo Red dye residues after the centrifugation (Figure 7.12 (2)). These results indicated that the cellulose production operon has been successfully built in *E. coli*.
7.3.5. SDS PAGE Analysis

As mentioned in previous chapter, BC is formed between the outer and cytoplasm membranes of the cell [37], consequently the recombinant proteins are located either in periplasm or in between membranes. Cmcax is secreted into growth medium by bacteria [47], while CcpAX colocalizes with BcsD [208], which is a periplasmic protein [209, 210]. BcsA is an integral membrane protein that contains the catalytic domain and forms a trans-membrane (TM) pore across the inner membrane. BcsB is a periplasmic protein which guides the polymer across the periplasm toward the outer membrane [210]. BcsC is a poreforming outer membrane protein.

*E. coli* HMS174 (DE3) (pBCS + pCMP) induced by 0.025 mM IPTG and the samples were collected 3 hours and 18 hours after induction at 37°C. The proteins extracted from induced cells were compared with proteins extracted from native strain (*E. coli* BL21 (DE3)). The SDS-PAGE gels are presented in Figure 6.19. Similar to study presented in previous chapter, total intracellular proteins and cytoplasmic fractions were also extracted to analyse by SDS analysis. The high protein concentrations in the extracts did not allow distinguishing the protein bands to observe any difference between protein extracts obtained from induced and un-induced cell groups (data not presented).
The sizes of CcpAx, Cmc, BcsA, BcsC, and BcsD are $39.0$, $37.2$, $80.0$, $89.6$, $144.8$, and $18.5$ kDa, respectively. In Figure 7.13 (A), Line 1 shows the molecular weight marker, while Line 2 shows periplasmic extracts of native cells. Lane 3 and Lane 4 shows the extracts of induced cells collected 3 hours and 18 hours after induction. In Figure 7.13 (B), Line 1 shows the molecular weight marker, while Line 2 shows insoluble fractions of native cells. Lane 3 and Lane 4 shows the extracts of induced cells collected 3 hours and 18 hours after induction. A band corresponding size of BcsC was detected in the insoluble fraction, which could be a combination of membrane proteins and inclusion bodies, while the other proteins were detected on periplasmic fraction. However CcpAX and Cmc were detected on the gel slightly upper than the region that correspond the sizes of these proteins. This ‘gel shifting’ could be originated from the additional sequences on cloning side, which creates differences in gel mobility of the proteins, as explained in Section 6.3.5 [211].

SDS-PAGE analysis of proteins extracted from *E. coli* BL21 (DE3) (pBCS + pCMP) exhibited thicker protein bands with an intensive colour, which indicates the higher expression of target genes (Chapter 6.3.5). Although the same concentration of total proteins were loaded to gel, the target proteins, extracted from *E. coli* HMS174 (DE3) (pBCS + pCMP) showed thinner bands as a consequence of low IPTG (0.025 mM) and low temperature (22˚C) used in fermentation. Especially intensity of the band corresponding to the size of BcsC in insoluble fraction is lower in this strain, which could be an indication of reduced accumulation of this protein.
7.3.6. Direct Detection of BC in Cultures

As bacterial cellulose is an extracellular material secreted directly into culture medium by bacteria, it was expected to be detected in the culture media of genetically modify stains. After 24 hours culturing, BC formation was clearly noticeable in the fermentation cultures of HMS174 (DE3) (pBCS + pCMP) (Figure 7.14 (2)). In contrast to this, no BC production was observed in the cultures of native HMS174 (DE3) cells, which were used as negative controls (Figure 7.14 (1)).

The BC produced by *E. coli* HMS174 (DE3) (pBCS + pCMP) was rather different than the BC produced by *A. xylinum*. Typically, BC produced by *A. xylinum* is generated a dense material, in shape of thick layer which covers the top of the culture (Figure 7.14, (3)) [224]. However, the product formed by *E. coli* HMS174 (DE3) (pBCS + pCMP) was in a form of dispersed pieces of fibres in the culture media, which is shown in Figure 7.14 (1). This could occur due to the difference between the swimming motility abilities of *E. coli* and *A. xylinum*. Active motility is dependent on a flagellar apparatus that is necessary for them to swim in liquid medium. It has been proved that motility itself was required for extracellular substance production the biofilm formation capacities directly correlated with
ability to swim. *E. coli* cells tend to settle in the bottom of cell culture in the absence of agitation. Therefore, they require high agitation speeds in liquid cultures, which could be have an effect on formation of the product. On the other hand, *A. xylinum* is able to produce cellulose even under static conditions.

**Figure 7.14:** Cultures of *E. coli* HMS174 (DE3) (1), *E. coli* HMS174 (DE3) (pBCS + pCMP) (2), and *A. xylinum* (3).

### 7.3.7. SEM Observation of BC

The morphology of the product was characterised using scanning electron microscopy (SEM) at an appropriate magnification. The shape and size distribution of the fibers was clearly demonstrated by SEM micrographs of the product.

SEM micrographs of BC produced in this study are shown in Figure 7.15 (2,3,4,5) and also a micrograph of BC produced by *A. xylinum* is also shown (Figure 7.15 (1) [225]. The diameter of BC produced by *A. xylinum* lies in nanoscale between several and 100 nm, while its length can be more than 100 μm. [6, 29, 226].
Figure 7.15: Scanning electron micrographs showing the microstructure of BC fibers produced by *A. xylanum* (1) and *E. coli* HMS174 (DE3) (pBCS + pCMP) (2,3,4,5)
In this study, the fibres produced by *E. coli* HMS174 (DE3)(pBCS + pCMP) revealed a comparable structure with BC cellulose produced by *A. xylinum*. The micrographs indicated that after 24 hours of cultivation, BC fibrils were very smooth and created a dense and random network, which shows a similar structure with BC produced by *A. xylinum*. In fact, the BC biosynthesised in this study exhibited remarkable structure. The lengths of the fibres were detected approximately 1000-1400 μm and diameters of the fibres are between 10–20 μm. The fibres were approximately 400 times longer and 200 times wider than BC cellulose produced by *A. xylinum*. In addition, SEM micrographs show the absence of biological impurities.

### 7.3.8. ATR-FTIR Characterisation of Biosynthesised BC

In order to obtain information about the purity of the bacterial cellulose produced ATR-FTIR spectra were recorded between 4000 and 600 cm\(^{-1}\), at room temperature (21±1°C), by a Spectrum FTIR Spectrometer (Perkin Elmer, USA).

Spectral band assignment of cellulose has been investigated extensively [227, 228]. A typical ATR-FTIR spectrum of BC reveals particular spectral bands at high wavenumbers. Table 7.4 summarised the main absorption bands and assignment for the ATR-FTIR spectra of bacterial cellulose.
Table 7.4: Main absorption bands and their assignments for the ATR-FTIR spectra of bacterial cellulose.

<table>
<thead>
<tr>
<th>Wavelength</th>
<th>Band Assignment</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>3500-3100</td>
<td>O-H stretching</td>
<td>[229]</td>
</tr>
<tr>
<td>2850-2950</td>
<td>C-H stretching</td>
<td>[230]</td>
</tr>
<tr>
<td>1635</td>
<td>-OH bending &amp; absorbed H₂O</td>
<td>[227]</td>
</tr>
<tr>
<td>1537</td>
<td>Amide I group</td>
<td>[230]</td>
</tr>
<tr>
<td>1385</td>
<td>C-H in-plane bending</td>
<td>[227]</td>
</tr>
<tr>
<td>1243-900</td>
<td>C-O-C stretching (glycosidic link)</td>
<td>[231, 232]</td>
</tr>
<tr>
<td></td>
<td>C-O stretching (C-3 &amp; C-6)</td>
<td>[227, 233]</td>
</tr>
</tbody>
</table>

The spectra are characterized by very strong absorption bands located between 900 and 1243 cm⁻¹ attributed to the C-O, C-O-C stretching vibration of glucose [227, 231-233]. The spectral bands observed at 1385 and 1420 are corresponding to the in-plane bending of the former C-H groups [227]. A noticeable region of the spectra exhibiting strong absorption bands appeared between 3100 and 3500 cm⁻¹ and corresponded to the stretching of –OH groups present in cellulose [229]. -OH bending and stretching modes appeared at 1635 cm⁻¹ to -OH bending of H₂O absorbed onto the BC network and also to -OH groups of BC itself. The analysis the spectra revealed that C-H stretching vibrations of BC (including -CH₂ and -CH₃) bands occurred between 2950 and 2850 cm⁻¹ [230]. Another noticeable region of the spectra appeared at 1537 cm⁻¹ which corresponds to amide I group [234].
Figure 7.16: ATR-FTIR spectra of biosynthesised BC by HMS174 (DE3) (pBCS + pCMP).

Figure 7.17: ATR-FTIR spectra of biosynthesised BC by A. xylinum [110].
Similar absorption bands were detected in ATR-FTIR data obtained from BC produced by *A. xylinum*, which was previously studied in our laboratory conditions (Figure 7.17). The band assigned to O-H stretching vibration of cellulose is identified between 3500 and 3100 cm⁻¹. The spectral band located between 2800 and 2900 cm⁻¹ is assigned to C-H stretching vibrations of BC (including -CH₂ and -CH₃), whereas the spectral bands located between 1420 and 1278 cm⁻¹, are corresponding to the in-plane bending of the former C-H groups. Furthermore, the band at 1640 cm⁻¹ is assigned to -OH bending. A stronger -OH bending was detected in BC produced by *E. coli*, which was possibly caused by higher water absorption in the sample compared to BC produced by *A. xylinum*.

### 7.4. Discussion

The host-vector interactions are extremely complex and unpredictable, which contributes to the stress level of the host. Expression of six recombinant proteins simultaneously under the control of strong T7lac promoter may require superior strains, which could maintain the plasmids inside. HMS174 (DE3) exhibits high plasmid stability as well as some target genes are more stable in HMS174 (DE3). In this chapter, BC production was investigated in genetically modified HMS174 (DE3), by transfection of recombinant plasmids pBCS and pCMP.

Fermentation of of *E. coli* HMS174 (DE3) (pBCS + pCMP) cells were performed at 37°C, 30°C and 22°C, which are also the same temperature used for the fermentation of *E. coli* BL21 (DE3) (pBCS + pCMP) in pervious chapter.

After the induction at 37°C, neither the cells induced by 0.025 mM IPTG nor the cells induced by 0.2 mM IPTG were able to reach 1.5 optical densities in 10 hours. As mentioned in Chapter 6, the reason behind the slow growth rate of genetically modified strains is that they experience additional burden caused by plasmid maintenance and replication [202]. The plasmid stability profile of the induced HMS174 (DE3) (pBCS + pCMP) cells exhibited a similar trend with that of BL21 (DE3) (pBCS + pCMP) and none of the genetically modified cells was plasmid-
bearing after 10 hours. On the other hand, un-induced HMS174 (DE3) (pBCS + pCMP) cells showed an improved plasmid stability profile (91%), which was only 60% in BL21 (DE3) (pBCS + pCMP). One of the reasons for plasmid instability is the use of elevated growth temperatures (higher than 30°C), a result of the metabolic stress are eclipsed by the toxicity presented by the production different proteins [203]. Also, it was reported that growth at 37°C caused some proteins to accumulate as inclusion bodies, which limited the final biomass. When the fermentation temperature was decreased to 30°C, remarkably, the final percentage of plasmid-bearing cells induced by 0.025 mM or 0.2 mM IPTG improved to 97% and 82%. The final optical density of the induced cultures also improved to 1.8 at 600 nm. However BC production was not detected at 30°C. That proved that although the plasmids are compatible with both BL21 (DE3) and HMS174 (DE3), stability of the expression system is stain-specific [122, 235].

It was reported that prolonged (e.g., overnight) induction at low temperatures may prove optimal for the yield of active proteins [207]. When the cells cultured at 22°C, finally BC production was detected after 24 hours. 86% of the induced cells were plasmid bearing after 10 hour and the growth profiles of induced and un-induced cells exhibited an extremely similar profile. The cells were cultivated in bioreactor under the same conditions (22°C, 0.025 mM IPTG). Although the plasmid stability of the cells cultured in bioreactor was more than 80%, the growth was not improved compared to that obtained in flasks. In addition, BC production was not obtained. The limited growth of stain HMS174 (DE3) possibly is a result of the O₂ supply problem, as reported also in previous chapter. Maintenance of O₂ at the set point could be a possible solution for the improvement of cell growth and also cellulose production. As a result of low O₂ levels, high acetate accumulation could occur in the culture. The acetate accumulation in HMS174 (DE3) cultures was detected approximately 5 times more than in the culture of BL21 (DE3)[236]. In order to maintain the pH in a certain range in bioreactor (7.0±0.25), a large volume of 4M NaOH was provided to media during the 24 cultivation, which caused a high nutrient dilution in media and this possibly limited the growth. BC production was not obtained in bioreactor, as it has
been proved that low O\textsubscript{2} concentration alters the metabolism of bacteria and also high salt concentration effects exopolysaccharide (EPS) and biofilm formation in bacteria negatively [237, 238].

Gene expression levels were also investigated in cellulose-producing HMS174 (DE3) (pBCS + pCMP) cells and the data revealed that all of the target genes expressed 3 hours after IPTG induction in low levels, which was between 1.4-2.7 fold. Similar to the gene expression data of BL21 (DE3), there was no significant difference between the expression levels of the genes at the same time point (P > 0.05). This proved that the expression levels of the genes are in stoichiometric balance. For effective expression of membrane proteins, transcription of the expression plasmids should be slowly induced with lower IPTG concentrations, to increase the solubility and activity of some target proteins [101, 214]. Moreover, SDS-PAGE analysis qualitatively demonstrated that there was less protein expression in HMS174 (DE3) (pBCS + pCMP) cells, which is more desirable for the expression of active proteins, in contrast to BL21 (DE3) (pBCS + pCMP) cells (Chapter 6.3.5). Cellulose production ability of the genetically modified HMS174 (DE3) stain was tested and confirmed by Congo binding analysis. All these findings supported the success of HMS174 (DE3) (pBCS + pCMP) in the production of BC.

Under aerobic respiration conditions, when the culture maintains a high growth rate during the exponential phase of growth, they secrete acetate into the media and this can inhibit cell growth or recombinant protein production. Marisch and colleagues compared the production of acetate K–12 and B strains, which was much higher for the K-12 strains compared to B strains. In this chapter, it was reported that of the HMS174 (DE3) cells (K-12 derived) exhibited a limited the growth compared to data reported on strain BL21 (DE3) in Chapter 6. High acetate accumulation may be one of the reasons for limited cell growth as well as the stress caused by high plasmid retaining capacity.

The structure of the material produced in this study was rather different and remarkable compared to cellulose produced by \textit{A. xylinum}. The fibres were approximately 400 times longer and 200 times wider than BC cellulose produced by
A. xylinum. This may be related with export mechanism and pore size in each bacterial strain. However there is still very limited information about the function and localization of export proteins. The diameter of the channels formed by E. coli for export of EPSs approximately 25-30 nm, while this is 12-15 nm for the export of cellulose in A. xylinum [239, 240].

The ATR-FTIR spectra of the BC samples revealed an identical chemical composition of the BC. However a noticeable region of the spectra exhibiting strong absorption appeared at 1537 cm\(^{-1}\), which corresponds to amide I group. Cellulose synthesis is correlated with biofilm formation and expression of multicellular behaviour (rdar morphotype) in E. coli [31, 116, 241]. As biofilm is an extracellular polymeric substance composed of extracellular DNA, proteins, and polysaccharides, existence of amide I group confirms presence of the proteins in the material.

The productivity of BC biosynthesized by the system reported in this chapter was extremely low (5 ±0.05 mg/L) at 22°C, in the presence of 0.025 mM IPTG. In effort to increase the productivity, new stains that can grow faster were investigated. BL21 (DE3) cells encountered plasmid loss but maintained growth [122]. Two mutant strains of BL21 (DE3), called C41 (DE3) and C43 (DE3) are frequently used to overcome the toxicity associated with overexpressing recombinant proteins using the bacteriophage T7 RNA polymerase expression system. When the toxicity of the plasmids is extremely high for replication in the strain BL21 (DE3), the toxic proteins can be expressed successfully in C41 (DE3). Dumon-Seignovert and colleagues investigated the ability of protein expression and stability of plasmids in BL21 (DE3) and C41 (DE3) [242]. After induction, the expression of heterologous proteins in mutant strain is generally better than in BL21 (DE3). The stability of a subset of the plasmids (3 h after induction) was determined in C41 (DE3) and varies from 62 to 92%. The study demonstrated the usefulness of strains C41 (DE3) in solving the problem of plasmid instability during the expression of toxic recombinant proteins. In addition OD\(_{600}\) of GM BL21 (DE3) cultures reached approximately 3.0, at 22 °C in the presence of 0.025 mM IPTG, whereas this was only 2.2 in the cultures of GM HMS174 (DE3) at the end of 10
hours culturing. Limited growth of GM HMS174 (DE3) leaded the further investigations to focus on a superior strain in membrane protein expression: *E. coli* C41 (DE3).

### 7.5. Conclusion

The work reported in this chapter proved that the type of strain, the fermentation conditions, and the concentration of inducer is high effective on system stability and productivity. Although BC production was achieved by strain HMS174 (DE3), the growth profiles showed a limited growth, which limits the yield of the product. Considering the fast growth rate and high plasmid stability of C41 (DE3) compared to K-12 derived strains, such as HMS174 (DE3), the further investigations focussed on C41 (DE3) for the investigation of BC production in *E. coli*. 
8. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED *E. coli* STRAIN 3: C41 (DE3)
8.1. Introduction

8.1.1. E. coli C41 (DE3) as a Host Microorganism

Although many membrane proteins can be overexpressed in inclusion bodies, their refolding into functional proteins is often not successful [243]. To solve the refolding problem, overexpression of membrane proteins by accumulation in the cytoplasmic membrane is required.

However, overexpression is often toxic to the cells, which inhibits biomass formation and severely reduces product yields [120]. For this reason, membrane protein overexpression should be optimized. Nevertheless no high-throughput and compatible method is available for the optimization process.

A decade ago, two new BL21 (DE3) strains have been developed, which allow high expression of a wide variety of toxic proteins previously difficult or even impossible to express in bacteria. The efficiency of these new strains, C41 (DE3) and C43 (DE3), in expressing toxic proteins has been proved in more than 350 publication [119, 120, 242]. Overexpression of many membrane proteins in these strains is hardly toxic, often resulting in high overexpression yields [119]. The reason for their improved membrane protein overexpression characteristics is still remains to be elucidated.

These strains comprise genetic mutations phenotypically designated for conferring toxicity tolerance [119, 242]. The strain C41 (DE3) was derived from BL21 (DE3) and it has at least one uncharacterized mutation that avoids cell death associated with expression of many toxic recombinant proteins. It was reported that the mutation could be related wit lacUV5 operon or T7 RNA polymerase expression [244]. As in standard BL21 (DE3) strains, C41 cells also carry the Lambda DE3 lysogen, which expresses T7 RNA polymerase from the lacUV5 promoter by IPTG induction. This allows the strain to express any gene cloned into a plasmid containing the T7 promoter.
Table 8.1: Comparison of C41 (DE3) and C43 (DE3) cells with the parental strain BL21 (DE3) in transformation and expression of heterologous proteins.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Transformation Success Rate</th>
<th>Expression-Induced Toxicity</th>
<th>Expressing Plasmids</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21 (DE3)</td>
<td>62%</td>
<td>96%</td>
<td>54%</td>
</tr>
<tr>
<td>C41 (DE3)</td>
<td>100%</td>
<td>50%</td>
<td>96%</td>
</tr>
<tr>
<td>C43 (DE3)</td>
<td>100%</td>
<td>4%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Dumon-Seignovert and colleagues investigated ability of transformation, expression the toxic proteins and stability of plasmids in BL21 (DE3) and C41 (DE3) [242]. According to their results of their study, summarised in Table 8.1, C41 (DE3) is the most advantageous strain for the expression of toxic membrane proteins with high plasmid stability and transformation success rate.

There are slight alterations in membrane protein overexpression kinetics between C41 (DE3) and its derivative C43 (DE3) [119]. Wagner et al. showed that, expression of LacY in C43 (DE3) is slightly delayed after IPTG induction. Consistently, the expression of T7RNA polymerase and the recombinant protein, and perturbation of the membrane proteome are slightly delayed in C43 (DE3) compared with C41 (DE3). They also observed relatively low expression levels of GFP and other soluble proteins in C43 (DE3). In conclusion, transcription kinetics in C43 (DE3) could not allow the overexpression of soluble proteins. For these reasons, C41 (DE3) selected as the host organism for the further investigation on BC biosynthesis in *E. coli.*
8.1.2. Aims and Objectives

The aim of the work reported in this chapter is to investigate BC by genetically modified strain of *E. coli* C41 (DE3) carrying two expression plasmids: pBCS and pCMP. The objectives of the work reported in this chapter are:

- Optimisation of bioprocess conditions (temperature, IPTG concentration) for maximum cell growth and plasmid stability.
- To verify the expression of target genes.
- To characterise the product by SEM and FTIR.

The presented chapter is composed of 5 parts: introduction, experimental methodology, results, discussion, and conclusion.

8.2. Methodology

The strains and recombinant plasmids used in this chapter are shown in the following Table 8.1. Other materials used in this chapter characterized in Chapter 4.1. The experiments reported in this chapter were conducted in order to investigate the effect of temperature and IPTG concentration on cell growth and plasmid stability of C41 (DE3) cells. For this purpose, the cells were cultured at 37°C and 180 rpm, in LB medium (described in Chapter 4.1.3) supplemented by streptomycin (50 mg/ml), carbenicillin (50mg/ml) and with 4 g/L glucose to prevent basal expression before IPTG induction. The experiments were performed in flasks unless otherwise noted. The media was supplemented by various concentrations of IPTG when the optical density reached 1.0 at 600 nm for induction of the expression of the target *bcs* genes. Following IPTG induction, media was supplemented with glycerol as carbon source culture temperature shifted to 30°C or 22°C, if necessary. After 24 hour culturing, cells were collected for further analysis. Samples were collected every 2 hours from media to monitor cell growth and plasmid stability as described in Chapter 4.2.6 and Chapter 4.2.5, respectively. Optical densities of cells were measured at 600 nm using spectrophotometer to
monitor cell growth and the samples were diluted to maintain the spectroscopic readings between reliable limits (0.2-0.8). The proportions of the cells maintain the target plasmid were determined by plasmid stability test and calculated in percentages. The colonies were counted to determine the colony forming units on each Petri dish. Average values from triplicates plates were taken. Proportions of the viable cells maintain the plasmids were calculated in percentages.

RNA extraction was performed for the confirmation of the recombinant gene expression. RNAprotect Bacteria Reagent was used to stabilize RNA is before bacterial cells are lysed in order to prevent the degradation of RNA transcripts. Following the cell lysis, the RNeasy Mini Kit or was used to purify total RNA. Details of the procedure can be seen in Chapter 4.2.11. Following that, Quantitech Reverse Transcription and QuantiTect SYBR Green PCR Kit kit were used for reverse transcription and Real Time q-PCR analysis, respectively. Details of the procedures can be seen in Chapter 4.2.15 and Chapter 4.2.16.

The functional groups in cellulose samples were identified by FTIR spectroscopy. ATR-IR spectra were recorded using a Spectrum One FTIR-spectrometer (Perkin Elmer, Massachusetts, USA). The spectra were collected at a resolution of 2 cm\(^{-1}\) in the range of 600 and 4,000 cm\(^{-1}\). The morphology and the microstructural features of samples were also investigated using a field emission scanning electron microscope (FESEM), LEO 1525 (Zeiss, Germany) operating at 5kV. Prior to analysis, samples were dried until constant weight and placed on a double-sided carbon tape mounted onto an aluminium stub. They were gold coated for 2 min at 20mA using an Emitech K575X Peltier (Ashford, UK) cooled sputter coater.
8.3. Results

8.3.1. Expression Systems

In this Chapter, the experiments were performed by *E. coli* C41 (DE3) host strain bearing the expression plasmids: pBCS carrying *bcsABCD* operon and pCMP carrying upper *bcs* operon (*cmcax* and *ccpAx*). The expression of the recombinant proteins was achieved by inducing the strong T7lac promoter by various concentrations of IPTG.

Table 8.2: *E. coli* strains and plasmid vectors used in Chapter 8.

<table>
<thead>
<tr>
<th><em>E. coli</em> Strain</th>
<th>Feature</th>
<th>Recombinant proteins synthesized by IPTG induction</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> C41 (DE3)</td>
<td>Native Strain</td>
<td>-</td>
</tr>
<tr>
<td><em>E. coli</em> C41 (DE3) (pBCS + pCMP)</td>
<td>Genetically Modified Strain carrying pBCS and pCMP</td>
<td>BcsA, BcsB, BcsC, BcsD, Cmcax, CcpAx</td>
</tr>
<tr>
<td><em>E. coli</em> C41 (DE3) (pBCS)</td>
<td>Genetically Modified Strain carrying pBCS</td>
<td>BcsA, BcsB, BcsC, BcsD,</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plasmids</th>
<th>Feature</th>
<th>Expressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBCS</td>
<td>pETDuet-1 with inserted <em>bcsABCD</em> operon encoding cellulose synthase</td>
<td><em>bcsA, bcsB, bcsC, bcsD</em>,</td>
</tr>
<tr>
<td>pCMP</td>
<td>pCDFDuet-1 with inserted <em>cmcax</em> and <em>ccpax</em> encoding endo-β-1,4-glucanase and cellulose complementing protein, respectively.</td>
<td><em>cmcax, ccpAx</em></td>
</tr>
</tbody>
</table>
In effort to investigate the effect of simultaneous expression of the BC biosynthesis proteins on *E. coli* C41 (DE3) (pBCS + pCMP), the cells were induced by various IPTG concentrations (0.025, 0.05, 0.2, 0.5, 1.0 mM) at various culture temperatures (22, 30, and 37 °C). *E. coli* C41 (DE3) cells, native strain, were used as control. Expression systems used in experiments and the proteins expressed using these systems are summarized in Table 8.2. In this chapter, *E. coli* C41 (DE3) (pBCS) cells also were cultured to investigate the BC production in the cells transformed by only pBCS.

### 8.3.2. Effect of IPTG Induction on Growth and Plasmid Stability of *E. coli* C41 (DE3)

In chapter 6 and 7, it was demonstrated that determination of optimum temperature and the concentration of inducer are critical for the system stability and the production.

The data reported in Chapter 7 showed that when *E. coli* HMS174 (DE3) (pBCS + pCMP) cells were induced by 0.025 mM IPTG at 22°C, cellulose production was detected. Consequently, cellulose biosynthesis in *E. coli* C41 (DE3) (pBCS + pCMP) cells was investigated under the same conditions (0.025 mM IPTG, 22°C) as a first step. Control groups used in these experiments are summarized in Table 8.3. Also, experimental conditions used in the work reported in this chapter were summarized in Table 8.4.

The fermentations were conducted in 500 ml flasks using 100 ml of LB culture unless otherwise noted. The experiments were performed in triplicate and the bars represent the range of error. Induction time points are indicated by arrows. Each fermentation was repeated three independent times and each analysis was replicated also three times, unless otherwise noted. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to $P < 0.05$. Glycerol was used as a main carbon source and the media supplemented 4 g/L of glucose to prevent basal expression that may cause extra burden to the cells in stationary phase.
When the OD value reached 1.0, *E. coli* C41 (DE3) (pBCS + pCMP) cells was induced by 0.025 or 0.2 mM IPTG. *E. coli* C41 (DE3) cells and un-induced *E. coli* C41 (DE3) (pBCS + pCMP) cells were used as controls.

**Table 8.3:** Control and experimental groups used in Chapter 8.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><em>E. coli</em> C41 (DE3)</td>
<td>Negative Control</td>
</tr>
<tr>
<td><em>E. coli</em> C41 (DE3) (pBCS + pCMP)</td>
<td>Un-induced Control</td>
</tr>
<tr>
<td><em>E. coli</em> C41 (DE3) (pBCS + pCMP) + IPTG</td>
<td>Induced Group</td>
</tr>
</tbody>
</table>

**Table 8.4:** Experimental conditions used in Chapter 8.

<table>
<thead>
<tr>
<th>IPTG (mM)</th>
<th>Temperature (°C)</th>
<th>Fermentation Strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>22°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>30°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.025</td>
<td>37°C</td>
<td>Flasks</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Growth profiles and plasmid stability profiles of C41 (DE3) cells cultured at 22 °C are shown in Figure 8.1 and 8.2, respectively. Before IPTG induction, the native cells showed a rapid growth and reached an optical density of approximately 1.61 at 600 nm. This was followed by un-induced cells (OD<sub>600</sub> = 1.11). After the
induction, unexpectedly, both induced groups reached slightly higher optical densities, approximately $OD_{600} = 2.50$, which could be an indicator of successful protein over-expression. On the other hand, native cells and un-induced cells reached $OD_{600} = 1.95$ and $OD_{600} = 1.92$, respectively. The growth profiles between induced and un-induced groups did not exhibit a significant difference ($P > 0.5$).

Before IPTG induction, approximately 100% of *E. coli* C41 (DE3) (pBCS + pCMP) cells were plasmid-bearing. 2 hours after IPTG induction, a slight decrease was detected in the proportion of both induced groups. After 24 hours culturing, the proportion of plasmid bearing cells among the cells induced by 0.025 mM was 96%, while this was 83% IPTG among the cells induced by 0.2 mM. This could explain the improvement in the cell growth of induced cells, since the faced less metabolic when the plasmid replication is lower.

**Figure 8.1:** Growth profiles of C41 (DE3) cells at 22°C.
After 10 hours culturing, agitation reduced to 90 rpm for further 14 hours fermentation. In the end of the fermentation period, a slight difference observed in the BC concentrations. The cells induced by 0.2 mM produced cellulose with a concentration of 8.5 mg/L, while the BC concentration of the cells induced by 0.025 mM IPTG was 13.5 mg/L.

![Figure 8.2: Proportion of the plasmid-bearing C41 (DE3) cells at 22 °C.](image)

In the next step, cultivation of the cells was performed at 30°C in effort to achieve a high density culture for high yields of production, since bacteria has shorter generation time at high temperatures. The influence of various IPTG concentrations (0.025 0.05 0.2 0.5 1.0 mM) was investigated at 30°C. According to data presented in Figure 8.3, the growth of the cells induced by 0.05 and 0.2 mM IPTG were not improved compared to those of cultured at 22°C. Unexpectedly,
after the induction, all induced groups reached the higher optical densities compared to native and un-induced control groups. The cells induced by 0.5 or 1.0 mM IPTG showed a significant increase (P < 0.05) and reached \( OD_{600} = 2.90 \) and \( OD_{600} = 2.64 \), respectively.

**Figure 8.3**: Growth profiles of C41 (DE3) cells at 30°C.

The percentages of plasmid-bearing cells were also monitored (Figure 8.4). After 24 hours cultivation, only 9% of the cells induced by 1.0 mM IPTG were plasmid bearing, while the percentage was 70% for the cells induced by 0.5 mM IPTG. This data demonstrated again that high IPTG concentrations contribute to the stress level of the host. Also, the unexpected improvement in the cell growth of the cell induced by 0.5 mM and 1.0 mM of IPTG could be explained by less metabolic
stress when there is no plasmid replication inside the cells. The product formation was not detected in these cultures. On the other hand, the groups induced by 0.05 and 0.2 mM IPTG exhibited high plasmid stability with 94% and 93%, respectively. The product formation was observed 5 hours after IPTG induction for both cultures. In the end of 24 hours cultivation, the concentration obtained by the cells induced by 0.2 mM IPTG was 21.0 mg/L, while the concentration was 31.1 mg/L in the cultures of the cells induced by 0.025 mM IPTG.

![Figure 8.4: Proportion of the plasmid-bearing C41 (DE3) cells at 30 °C.](image)

Cultivation of the cells was performed also at 37°C. Since the cultures induced by higher than 0.2 mM IPTG exhibited plasmid instability at 30°C, the cells induced by 0.025 and 0.2 mM IPTG. The cells showed similar growth profiles with
each other \((P > 0.05)\) without any significant improvement compared to the growth profiles of cells cultured at 22°C and 30°C, as presented in Figure 8.5.

![Growth profiles of C41 (DE3) cells at 37°C.](image)

**Figure 8.5:** Growth profiles of C41 (DE3) cells at 37°C.

Plasmid stabilities of the cells cultured at 37°C are shown in Figure 8.6, which exhibited a very high stability level. Although the final percentage of plasmid-bearing cells induced by 0.025 mM IPTG and 0.2 mM IPTG were detected as 98% and 93%, respectively. However, BC production was not detected under these conditions. This could be explained by the increase in the production of different metabolic proteins at high temperatures, which eclipses the effects of the metabolic burden.
Although it has been proved that \textit{ccpA}x and \textit{cmcx} are also essential genes for the production of BC, in order to simplify the system and reduce the metabolic stress on the cells, \textit{E. coli} C41 (DE3) (pBCS) cells, which was transformed by only pBCS, were cultured at 30° and induced by 0.025 mM IPTG. Although under these conditions production was detected in the cells harbouring both plasmids, the production was not obtained either of the two independent experiments conducted by the fermentation of the cells harbouring only pBCS. (Data not presented).

\textbf{Figure 8.6:} Proportion of the plasmid-bearing C41 (DE3) cells at 37 °C.

\textbf{8.3.3. Single plasmid approach}

Although it has been proved that \textit{ccpA}x and \textit{cmcx} are also essential genes for the production of BC, in order to simplify the system and reduce the metabolic stress on the cells, \textit{E. coli} C41 (DE3) (pBCS) cells, which was transformed by only pBCS, were cultured at 30° and induced by 0.025 mM IPTG. Although under these conditions production was detected in the cells harbouring both plasmids, the production was not obtained either of the two independent experiments conducted by the fermentation of the cells harbouring only pBCS. (Data not presented).
8.3.4. Gene Expression Analysis of BC Synthase Related Genes

q-PCR analysis was performed in order to confirm expression of the recombinant genes (*cmcax*, *ccpAx*, *bcsA*, *bcsB*, *bcsC* and *bcsD*) cellulose producing C41 (DE3) (pBCS + pCMP) cells. Housekeeping gene used in this analysis was *serC*, a chromosomal *E. coli* gene that is usually constitutively expressed [106]. RNA quantification was performed to balance the amount of RNA for each sample before the production of cDNA by reverse transcription.

The cells were cultured at 30°C and harvested for RNA extraction 3 and 18 hours after 0.05 mM or 1.0 IPTG induction. The relative gene expressions were showed in Figure 8.7. and Figure 8.8. Statistical analysis was performed using one-way analysis of variance (ANOVA) and statistically significance was assigned to P < 0.05.

![Relative Quantification](image)

**Figure 8.7:** Relative gene expressions of C41 (DE3) (pBCS + pCMP), 3, 6 and 18 hr after induction by 0.05 mM IPTG at 30°C.
3 hours after 0.05 mM IPTG induction, the relative quantification demonstrated that the only genes expressed were *cmcax* (2.01 fold) and *ccpAx* (2.12 fold), which were carried by pCMP (Figure 8.7). After that, expression levels slightly increased in 18 hours. The genes carried by pBCS were detected 6 hours after induction, between 1.13-1.46 fold. The expression levels of all genes increased after 18 hour induction. *bcs* genes expressed between 1.61-2.19 fold, while upstream genes *cmcax* and *ccpAx* expressed 5.02 and 3.32 fold, respectively. Although the data demonstrated that the relative expression of the *bcs* genes carried by pBCS was lower than the upstream genes carried by pCMP, all genes were expressed at significant levels in the end of culturing. (P < 0.05)

![Bar chart](image)

**Figure 8.8:** Relative gene expressions of C41 (DE3) (pBCS + pCMP) 3, 6 and 18 hr after induction by 1.0 mM IPTG at 30°C.
In contrast, the expression levels exhibited a decreasing trend when the cells induced by 1.0 mM IPTG (Figure 8.8). All target genes expressed at higher in the presence of 1.0 mM IPTG in comparison to the levels obtained in the presence of 0.05 mM IPTG. Expression levels increased significantly even after 3 hours of the induction (2.54 - 11.64 folds). The levels decreased slightly after 18 hours, which were detected between 1.49 - 7.11 fold.

The data revealed that, transcription of the expression plasmids was slowly induced with lower IPTG concentrations. Slow induction allows an effective expression of membrane proteins to increase the solubility and activity of some target proteins [101, 214]. This could be a possible reason for the detection of BC production in the cultures induced by low IPTG concentrations.

8.3.5. Direct Detection of BC in Cultures

BC is an extracellular substance secreted directly into culture media by bacteria. After 5 hours culturing at 22°C or 30°C, the product was clearly noticeable in the fermentation cultures of C41 (DE3) (pBCS + pCMP) (Figure 8.9 (2)). In contrast to this, no BC production was observed in the cultures of native C41 (DE3) cells, which were used as negative controls (Figure 8.9 (1)).

The BC produced by C41 (DE3) (pBCS + pCMP) exhibited a different morphology compared to BC produced by *A. xylinum*. As mentioned in Chapter 7.3.6, typically, BC produced by *A. xylinum* is a dense material, in shape of thick layer which covers the top of the culture (Figure 8.9, (3))[224]. On the other hand, the product formed by C41 (DE3) (pBCS + pCMP) was very similar with that of HMS174 (DE3) (pBCS + pCMP), in a form of dispersed pieces of fibres in the culture media, which is shown in previous chapter, Figure 7.14 (2).
Figure 8.9: Cultures of native strain (control) (1) and genetically modified strain (experiment) (2).

8.3.6. SEM Observation of BC

The morphology of the product was characterised using scanning electron microscopy (SEM) at an appropriate magnification.

The product displayed an irregular pattern compared to bacterial cellulose produced by *A. xylinum*, which exhibits sharp and distinct features of fibers, as presented in previous chapter (Figure 7.15 (1)). This discrepancy was attributed to biofilm formation, which is correlated with BC production. Morphology of biofilms exhibits a particular type of cell-cell interaction, which is rigid, but fragile connections between cells. [219]. The typical morphology of a biofilm produced by *E. coli* O157:H7, which is a pathogenic strain, is shown in Figure 8.10 (5,6) [245, 246]. The material produced by C41 (DE3) (pBCS + pCMP) displayed a similar structure to biofilms produced by pathogenic *E. coli* strains. The fibres with the length of approximately 600 μm were detected, which were embedded into the material (Figure 8.10 (3)). The size of the fibers was much bigger to BC produced by *A. xylinum*. Similar observations were detected in previous chapter in terms size of fibres produced by HMS (DE3) (pBCS + pCMP). Nonetheless, the overall structure was different from that reported in Chapter 7, where it was observed distinct features of fibers without any other substances.
Figure 8.10: Scanning electron micrographs showing the microstructure of BC produced by *E. coli* C41 (DE3) (pBCS + pCMP) after 24 hr fermentation (1,2,3,4), *E. coli* O157:H7 (5), *E. coli* PHL644 (6). Image 5 and Image 6 are adopted from Lee et al. and Tsoligkas et al., respectively [245, 246].
8.3.7. ATR-FTIR Characterisation of Biosynthesised BC

In order to obtain information about the purity of the bacterial cellulose produced ATR-FTIR spectra were recorded between 4000 and 600 cm\(^{-1}\), at room temperature (21±1°C), on a Spectrum FTIR Spectrometer (Perkin Elmer, USA).

Spectral band assignment of cellulose has been investigated extensively [227, 228]. A typical ATR-FTIR spectrum of BC reveals particular spectral bands at high wavenumbers. In previous chapter, the main absorption bands and assignment for the ATR-FTIR spectra of bacterial cellulose are summarised in Table 7.3.

The spectra of the material are characterized by very strong absorption bands located between 900 and 1243 cm\(^{-1}\) attributed to the C-O, C-O-C stretching vibration of glucose [227, 231-233]. The spectral band observed at 1386 is corresponding to the in-plane bending of the former C-H groups [227]. A noticeable region of the spectra exhibiting strong absorption bands appeared between 3100 and 3500 cm\(^{-1}\) and corresponded to the stretching of –OH groups present in cellulose [229]. The water bending and stretching modes appeared at 1631 cm\(^{-1}\) and corresponded to water absorbed from cellulose. The analysis the spectra revealed that C-H stretching vibrations of BC (including -CH2 and -CH3) bands occurred between 2950 and 2850 cm\(^{-1}\) [230]. Another noticeable region of the spectra appeared at 1516 cm\(^{-1}\), which corresponds to amide I group [234].
8.4. Discussion

The host-vector interactions are extremely complex and as a consequence optimisation of protein expression is tricky in a recombinant expression system. In Chapter 6, although stain BL21 (DE3) is a widely used for the over expression of the target proteins, it was concluded C41 (DE3) exhibits high plasmid instability. In Chapter 7, although strain HMS174 (DE3) showed high plasmid stability, cell growth and BC productivity were not satisfactory. In this chapter, BC production was investigated in genetically modified HMS174 (DE3), by transfection of recombinant plasmids pBCS and pCMP.

Fermentation of *E. coli* C41 (DE3) (pBCS + pCMP) cells were performed at 22°C, 30°C and, 37°C, which are also the same temperature used for the fermentation of *E. coli* in Chapter 6 and Chapter 7.

Strain C41 (DE3) is generally superior to its parent strain, BL21 (DE3), as a host for protein over-expression. It allows the expression of toxic membrane
proteins. In this chapter, this ability of strain C41 (DE3) was proved by broth and plasmid stability profiles. It was shown that after the induction at 22°C, both induced groups (0.025mM or 0.2 mM IPTG) reached higher optical densities, approximately OD$_{600}$ = 2.50. A similar observation was also made in the cells cultured at 30°C. The induced cells (0.5 or 1.0 mM IPTG) reached higher optical densities (OD$_{600}$ = 2.64-2.90) compared to native and un-induced cells, which could be result of less metabolic stress after plasmid loss. This indicated a low or no metabolic stress on the cells, in contrast to data reported on parental strain BL21 (DE3) and HMS174 (DE3) in previous chapters. The other potential reason may be the rapid growth of the native and un-induced control groups in the beginning of the cultivation, which possibly results in the accumulation of high concentrations of toxic materials such as acetate and this limits their growth compared to induced groups.

Dumon-Seignovert and colleagues investigated the plasmid stability in C41 (DE3) [242]. The stability of a subset of the plasmids (3 h after induction) was determined in C41 (DE3) and varies from 62 to 92%. In this thesis, more than 80% of the induced C41 (DE3) (pBCS + pCMP) cells were plasmid bearing in the end of 24 hours cultivation at 22°C or 30°C. A similar data was observed in HMS (DE3) (pBCS + pCMP) cells cultured at 22°C or 30°C. On the other hand, induced C41 (DE3) (pBCS + pCMP) cells also exhibited high stability event at 37°C in contrast to HMS (DE3) (pBCS + pCMP) and BL21 (DE3) (pBCS + pCMP). However, BC production was not detected at 37°C. Zhou et al. reported that prolonged (e.g., overnight) induction at low temperatures may prove optimal for the yield of active proteins [207]. This hypothesis was tested and the cells cultured at 22°C or 30°C, which resulted in BC production detected after 5 hours. Nevertheless, there was no significant difference in productivities. The only instability in plasmids was detected in C41 (DE3) (pBCS + pCMP), when the cells was induced by higher than 0.5 mM IPTG. Also, BC production was not detected in the cultures of C41 (DE3) (pBCS), which proves that $ccpAx$ and $cmcax$ are necessary for the BC biosynthesis.
Gene expression levels were also investigated in C41 (DE3) (pBCS + pCMP) (pBCS + pCMP) cells and the data revealed that all of the target genes expressed 6 hours after 0.05 mM IPTG induction and showed an increasing trend, which is an indication of slow induction. For effective expression of membrane proteins, transcription of the expression plasmids should be slowly induced with lower IPTG concentrations, to increase the solubility and activity of some target proteins [101, 214]. On the other hand, C41 (DE3) (pBCS + pCMP) (pBCS + pCMP) cells induced by 1.0 mM IPTG exhibited higher gene expressions, with an increasing trend, however, this resulted in no BC production.

The structure of the material produced in this study was different compared to the material produced by HMS174 (DE3) and A. xylinum. The fibers were sparse and embedded into the material. Zogaj et al. demonstrated that cellulose biosynthesis in E. coli occurs with the production of thin aggregative fimbriae [116]. The structure of the material produced in this chapter exhibited a biofilm-like material. The ATR-FTIR spectra of the samples revealed also supportive results for this theory. The ATR-FTIR spectra of the BC samples revealed an identical chemical composition of the BC as well as a noticeable region which corresponds to amide I group. As biofilm is an extracellular polymeric substance composed of extracellular DNA, proteins, and polysaccharides, existence of amide I group confirms presence of the proteins in the material.

Although initially biofilms were considered as undesired products, later on, there have been many investigations into the use of biofilms for useful purposes such as energy generation, wastewater treatment and biotransformations [246-249]. These studies have mainly focused on altering the population of the biofilm rather than the extracellular matrix itself. It has been described the design of synthetic genetic circuits, which control the population balance in a dual-species biofilm to control biofilm biosynthesis and dispersion based on quorum-sensing [250]. A novel protein engineering system, called Biofilm-Integrated Nanofiber Display (BIND), has developed based on curli formation in another recent impressive work. Using BIND, it has been demonstrated that bacteria could be
programmed to provide desired functions for the large-scale production of biomaterials [251]. Since biofilms are mostly produced by pathogenic and commensal *E. coli* strains, the genetically modified C41 (DE3) strain could have a great potential as a non-pathogenic strain for the use of biofilms for useful purposes.

### 8.5. Conclusion

Although the concentration of BC biosynthesized by the system reported in this chapter increased slightly up to 31.1 mg/L, the amount of the product was not satisfactory. Genetically modified C41 (DE3) strain exhibited higher plasmid stability even at 30 °C, and the product formation was obtained only 3-5 hours after induction. Therefore, this strain could be promising for design of a high density fed-batch culture. High cell density cultivation (HCDC) is a powerful approach developed to improve productivity and also to provide advantages such as reduced culture volume, lower production costs[77]. Fed-batch processes have widely used to obtain high cell-density. HCDC allows reaching a higher dry cell weight, consequently, a higher product concentration which is not possible in conventional batch cultures[252]. For this reason, further investigations focussed on design of a fed batch system with the aim of increasing the productivity.
9. INVESTIGATION OF BACTERIAL CELLULOSE PRODUCTION BY GENETICALLY MODIFIED *E. coli* C41 (DE3) IN FED-BATCH SYSTEMS
9.1. Introduction

9.1.1. Limiting Factors in Batch Cultivation

9.1.1.1. Carbon Source and Carbon catabolic repression in *Escherichia coli* strains

*E. coli* is capable of utilizing several compounds as carbon sources. However, glucose is the most widely utilized carbon source since it is easily to be metabolised as a monosaccharide. Although glycerol is an energy-poor carbon source, it has an enhanced biotechnological significance as carbon source since it is a by-product of the biodiesel synthesis, whose production is expected to increase in the future [253, 254].

Various carbon sources have been published that are suitable for recombinant protein expression in *E. coli*, including glucose, acetate, glycerol, galactose, mannitol, fructose, xylose and succinate [255-257]. Zhang *et al.* investigated the effect of different carbon sources on the expression of 6-O-sulfotransferase, which is a transmembrane protein required for the production of heparin. Glycerol, glucose and fructose were used to examine the cell growth, enzyme production and enzyme activity. In the presence of glycerol, the highest cell growth and highest active enzyme production were obtained [258].

The presence of a carbon source in the medium can repress expression of certain genes and operons. In the presence of glucose, the level of cAMP, which is a component directly stimulates transcription of RNA polymerase, are low. Consequently transcription from the *lac* promoter is also low. In order to avoid increased basal expression in DE3 lysogens, LB medium can be supplemented up to 10 g/L glucose. For this reason glycerol was used as a main carbon source in this study and the media supplemented 4 g/L of glucose to prevent basal expression that may cause extra burden to the cells. As glucose is the most easily metabolized monosaccharide, BC production in the present of glucose was investigated as well as glycerol. Glucose concentration was kept lower than upper limit (10 g/L) to prevent the catabolite repression after induction.
Nutrient exhaustion is a limiting factor for \textit{E. coli} growth. During the stationary phase, nutrient becomes exhausted, and toxic metabolites gradually accumulate. This could become a stress to the cell, and eventually leads to the death phase. Therefore, a feeding strategy should be obtained to provide nutrients for cells to generate high density cultures.

\subsection*{9.1.1.2. Acetate production and pH}

One of the key problems preventing the achievement of high volumetric productivity during fermentation of \textit{E. coli} is the conversion of an important portion of the commonly used carbon source glucose into acetate [259, 260]. Although glucose is the most preferred carbon source for \textit{E. coli}, acetate accumulation can cause negative effects on cell growth. Glucose is first converted into pyruvate by glycolysis and then to acetyl-CoA. Once acetyl-CoA is formed, it can be used in the tricarboxylic acid cycle (TCA) cycle in aerobic respiration to produce energy, which usually results in acetate accumulation, or it can be used in glyoxylate cycle (Figure 9.1).

Investigations on cultivation conditions, medium compositions, and genetic approaches have been used in order to minimize acetate formation [261]. The use of different carbon sources such as glycerol has been shown to reduce acetate formation and also improve the activity of the products [262, 263]. When \textit{E. coli} is cultured aerobically in the media supplemented with glycerol, this carbon source is incorporated into the central metabolism as dihydroxyacetone phosphate (DHAP), which can take part in both gluconeogenic and glycolytic processes [264]. In consequence, specific growth rate decreases and acetate production is produced in a low level [265].
Figure 9.1: The glucose (1) and acetate (2) metabolism in E. coli [266, 267].
Acetate formation can be avoided by using glycerol in the place of glucose, however, significant acetate elimination by recombinant *E. coli* grown on glycerol has also been reported, as well as a decrease in target gene expression [268, 269]. Wagnes *et al.* reported that replacing glucose with glycerol resulted in high plasmid instability in *S. aureus* protein, a recombinant expression system the under the lac promoter control hosted by *E. coli* [270]. It was concluded that glycerol grown cells are not capable of resisting the additional metabolic burden caused by the expression of recombinant proteins.

**Table 9.1:** pH ranges for cell growth, plasmid production and protein production.

<table>
<thead>
<tr>
<th>pH range</th>
<th>pH&lt;5.5</th>
<th>pH 5.5 to 8.5</th>
<th>pH&gt;8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell growth</td>
<td>Slow to stop</td>
<td>Optimal</td>
<td>Stop to cell death</td>
</tr>
<tr>
<td>Plasmid production</td>
<td>Suitable for most plasmids</td>
<td>Optimal</td>
<td>Suitable for some plasmids</td>
</tr>
<tr>
<td>Protein production</td>
<td>Optimal for some proteins</td>
<td>Optimal for most proteins</td>
<td>Poor for most proteins</td>
</tr>
</tbody>
</table>

pH is another significant restrictive factor for cell growth, plasmid production and protein production in addition to nutrient exhaustion and accumulation of acetate. Production of acetic acid during fermentation process lowers the pH of a medium. *E. coli* cells can grow over a range of three pH units (from pH 5.5 to 8.5)[271]. Extreme pH beyond this range significantly decreases the cell growth and even causes cell death (Table 9.1). pH has also significant effects on the activity of proteins. Specifically, BcsAB complex (cellulose synthase) exhibits a maximum catalytic activity at neutral pH, as presented in Figure 9.2 [272].
When the number of cells increases in a culture, oxygen availability becomes a limiting factor with significant influence on growth [273, 274]. This limitation activates the expression of more than 200 genes in order to adjust the metabolic capacities of the cell to the availability of oxygen, which decreases the growth rate [275]. Increasing the shaking speed could be a possible solution to increase the amount of available oxygen in shake flasks. However, specifically in this study, this could negatively affect the structure of the produced cellulose.

In addition to acetate, the accumulation of other intermediate metabolites to toxic levels in bacteria during a fermentation process causes a stress response. Because, oxygen limitation triggers the formation of numerous metabolites of the acid metabolism, which are undesirable, such as succinate, lactate, ethanol, and hydrogen [276, 277].

### 9.1.2. Fed-batch Cultivation and Mathematical Models

Fed-batch fermentation processes are frequently used in order to prevent the accumulation of toxic substrates or products consequently allowing the significant increases in cell density achievement of higher product concentrations [278, 279].

---

**Figure 9.2:** Activity of BcsAB complex in a pH range from 4.5 to 9.5 [271].
The acetate accumulation can be decreased by reducing the growth rate if the carbon source is supplied slowly. It was suggested that all the acetyl CoA produced by the glycolytic pathway can be consumed by TCA cycle handle and this abolishes the acetate formation [280].

During a fed-batch *E. coli* fermentation process, the system states change significantly, from a low initial biomass to very high biomass and product concentrations. This dynamic behaviour requires the development of optimisation methods to find the optimal strategy during the feeding stage in order to improve the process. The typical input variable in these processes is the inflow rate of carbon source as a function of time.

Several fed batch methods were developed for carbon source feeding to maintain acetate levels low. These methods are developed based on mathematical models that describe growth and death rate, the expected demand for nutrients and growth inhibition by acetate production. An optimal concentration of all components and feed rates can be calculated by using a dynamic model. Classic equations to determine the desired flow rate was proposed by Lee.[77]

Numerous mathematical models have been proposed for recombinant fermentations. However, only a few of them have been used to optimize industrial fermentations, due to the complexity of many of the models as well as the absence of important factors such as acetate inhibition.

### 9.1.3. Aims and Objectives

The aim of the work reported in this chapter is to design a fed-batch system in effort to improve the productivity and to overcome the limiting factors for bacterial cellulose production by genetically modified strain of *E. coli* C41 (DE3) carrying two expression plasmids: pBCS and pCMP. The objectives of the work reported in this chapter are:

- To monitor the effect of different carbon sources and concentrations on cell growth.
To monitor the consumption of carbon sources.

To monitor the effect of different carbon sources and concentrations on acetate production and pH.

To develop a dynamic model for the batch cultures.

To optimise BC production in fed-batch cultures based on developed dynamic model.

To characterise the morphology of the product by a light microscope.

The presented chapter is composed of 5 parts: introduction, experimental methodology, results, discussion and, conclusion.

9.2. Methodology

The strain *E. coli* C41 (DE3) (pBCS + pCMP) was utilised in this chapter to investigate BC production. Other materials used in this chapter characterized in Chapter 4.1.

Effect of different carbon sources at various concentrations on cell growth, acetate production, pH, and cellulose productivity was investigated for the selection of optimum conditions. The cells were cultured either in flasks or in bioreactor. LB medium was supplemented with glycerol or glucose as carbon source. The cells were induced by 0.05 mM IPTG for induction of the expression of the target *bcs* genes when the optical density reached 1.0 at 600 nm. After 24 hour culturing, cells were collected for further analysis.

Optical densities of cells were measured at 600 nm using spectrophotometer to monitor cell growth every 2 hr. Samples were diluted to keep the spectroscopic readings between reliable limits (0.2-0.8).

The rates of glucose and glycerol consumption were monitored to develop a feeding strategy. For this purpose, Glucose Assay Reagent (Sigma) or Free Glycerol
Reagent (Sigma) was used to detect the concentration of free glucose/glycerol amount in LB media. The samples were collected every two hours and centrifuged at 10000 rpm for 10 minutes to separate the cells from medium. Following the sterilization of supernatant by filter, they stored for free glucose determination assay. The composition of reaction mixture summarized Table 4.14. The absorbance was detected by UV spectrophotometer at 330 nm.

Free acetate concentration was also monitored to determine if the concentration reached the toxic levels by acetate determination kit (Sigma), which was explained in Chapter 4.2.19.3. Acetate and acetate substrate mix converted to an intermediate in the reaction catalysed by Acetate Enzyme Mix. The intermediate reduces the probe to a yellow-colored product and absorbance was read at 450 nm.

The morphology of the product was determined by light microscopy in this chapter at appropriate magnifications.

9.3. Results

9.3.1. Comparison of Growth, Carbon Source Consumption, Acetate Production of E. coli in Batch Cultures

The growth characteristics, carbon source consumption and acetate production of E. coli C41 (DE3) (pBCS + pCMP) were compared in the presence of various concentrations of glucose or/and glycerol to assess the potential of the strain to grow at high cell densities with less acetate production. In addition, a dynamic model was developed based on this preliminary experimental data for to determine the feeding strategy.

The cells were cultured at 30 °C and induced by 0.05 mM IPTG in LB media supplemented with glycerol or/and glucose. Experiments were conducted in flasks at 180 rpm and samples were collected every 2 hours. Experiments conducted in duplicates and the averages were presented.
First, the standard culturing procedure used in previous chapters was tested. LB medium was supplemented with 4 g/L glucose to prevent the catabolic repression in stationary phase. Following that, the media was supplemented by glycerol (4 g/L) at induction time and glycerol concentration was monitored. For the other cultures, LB media was supplemented with one of the following carbon source in the beginning of cultivation: 4 g/L glycerol, 4 g/L glucose and 10 g/L glucose.

Growth characteristics of the cells cultured in two different carbon sources are shown in Figure 9.3. The cultures supplemented with 10 g/L glucose reached the highest specific growth rate with 0.42 hr\(^{-1}\), which was followed by the cultures supplemented by 4 g/L glycerol with 0.33 hr\(^{-1}\). The specific growth rate of the cultures supplemented with two carbon sources and cultures supplemented by 4 g/L glucose were 0.28 hr\(^{-1}\) and 0.27 hr\(^{-1}\), respectively. As the glucose is an easily metabolised monosaccharide, in the presence of high concentration of glucose, cells reached the highest optical density with 2.72 at 600 nm as well as highest specific growth rate.
Consumption of carbon sources versus time is presented in Figure 9.4. When the media supplemented with 4 g/L glycerol or 4 g/L glucose, the consumption characteristics exhibited very similar profiles and both consumed completely in 24 hours. However, when the media supplemented with 10 g/L glucose, only 3.8 g/L of glucose was consumed. Despite the abundance of glucose in LB media, the cells stopped growing and entered to the stationary phase after 12 hours cultivation (Figure 9.1). As no nutrient exhaustion was detected, this was attributed to low pH of the culture due to the acetate accumulation, which inhibited the cell growth. The final glycerol concentration in the cultures supplemented with 4 g/L glycerol and 4 g/L glucose was 2.7 g/L. Glycerol was still abundant in the cultures due to the glucose supplement in the beginning of cultivation to prevent the catabolic repression. As glucose is easily consumed carbon source by cells, they tend to consume glucose first.
Figure 9.4: Consumption of carbon sources C41 (DE3) cells in LB media supplemented by two different carbon sources.

Production of acetate versus time is presented in Figure 9.5. The cultures supplemented with 4 g/L glucose and the cultures supplemented with 10 g/L glucose produced the highest amounts of acetate, 2.1 g/L and 2.3 g/L respectively. When the cultures supplemented with 4 g/L glycerol, the final concentration of acetate was lower, 1.6 g/L. The data revealed that the use of glycerol in replacement of glucose reduces the acetate formation. This has been also proved in literature [262, 263].
The cultures supplemented with both carbon sources produced slightly higher acetate, 1.8 g/L, compared to the cells supplemented with only glycerol, suggesting that part of the acetate produced when growing on glucose as the initial carbon source. The literature observations showed that growth stops when acetate levels reach 10-20 g/L [281-283]. However, tolerant to acetate inhibition could depend on especially host/vector used in a system. Acetate concentrations as low as 0.5 g/L have been shown to inhibit cell growth by 50% [284].

pH levels were in an agreement with acetate production profiles, as presented in Figure 9.6. pH profiles of the cultures showed similar trends except the cultures supplemented with 4 g/L glycerol, which exhibited the lowest acetate production especially in the early stages of culturing (0–8 hours). In consequence, they resulted in the highest final pH with 4.7. The cultures supplemented with 10
g/L glucose exhibited the highest acetate accumulation, which resulted in the lowest final pH with 4.3. When the acetate levels reached above 0.5 g/L after approximately 12 hours culturing, pH was below 5.0 in all cultures, which could be one of your reasons for the inhibition of the cell growth. One of the other reasons for the low pH of the culture could be the accumulation of other toxic intermediate metabolites formed during the acid metabolism in addition to acetate, as the acetate accumulation did not reached the limits that stops the cell growth.

Figure 9.6: pH profile of C41 (DE3) cells in LB media supplemented by different carbon sources.

Comparison of the growth characteristics of the cultures supplemented with different carbon sources were summarised in Table 9.2. Specific growth rates, consumption rate of glycerol/glucose, production rate of acetate were calculated for exponential growth phase of the cultures. Dry cell weights (DCW) were calculated by calibration chart, which was prepared by plotting optical density vs. cell dry weight as shown in Appendix 2. An equation as \( y = 2.2826x \) was obtained and the slope of this line was used to determine the dry cell weight of a known
Following the calculation total biomass, yield of biomass and acetate were calculated based on the amount of the total carbon source consumed. The yield of cellulose was calculated also depending on the amount of the total carbon source consumed.

The cultures supplemented with 10 g/L glucose exhibited the highest specific growth rate, 0.42 h\(^{-1}\), however it caused a higher acetate production rate as a consequence, which was calculated 0.112 g/L.h. and also resulted in 0.60 g/g acetate yield (grams of acetate per gram of carbon source utilised). The acetate production rates in the other cultures were similar, which were observed between 0.059-0.063 g/L.h. On the other hand, the rates of acetate production and glucose consumption were considerably different, which was unexpected. It has been proposed that glucose could be metabolised by glyoxylate cycle instead of TCA cycle [118, 285, 286]. The yield of biomass was observed between 0.21-0.32 g/g (grams of DCW per gram of carbon source utilised in the cultures). Although the yield of acetate is the lowest in the cultures supplemented with 4 g/L glycerol, the yield of cellulose observed lowest with 2.95 mg/g (milligrams of cellulose per gram of carbon source utilised in the cultures). Overall, the cultures supplemented with 4 g/L glucose and the cultures supplemented with 4 g/L glycerol exhibited very similar characteristics. However highest final cellulose yield was detected in the cultures supplemented with 4 g/L glucose with a yield of 22.50 mg/L. Using glucose as main carbon source could be more advantageous, as it allows preventing increased basal expression. For this reason, glucose was used for the further investigations on the development of fed-batch system.
Table 9.2: Comparison of growth characteristics of C41 (DE3) cells in LB media supplemented by different carbon sources.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Specific growth rate (1/h)</th>
<th>Glucose consumption rate (g/L.h)</th>
<th>Glycerol consumption rate (g/L.h)</th>
<th>Acetate production rate (g/L.h)</th>
<th>Final DCW (g/L)</th>
<th>Yield of acetate from carbon source (g/g)</th>
<th>Yield of acetate from carbon source (g/L.h)</th>
<th>Final cellulose concentration (mg/L)</th>
<th>Volumetric productivity (mg/L.h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LB + glucose (4 g/L)</td>
<td>0.27</td>
<td>0.32</td>
<td>-</td>
<td>0.062</td>
<td>1.11</td>
<td>0.27</td>
<td>0.50</td>
<td>5.50</td>
<td>22.50</td>
</tr>
<tr>
<td>LB + glucose (10 g/L)</td>
<td>0.42</td>
<td>0.36</td>
<td>-</td>
<td>0.112</td>
<td>1.23</td>
<td>0.32</td>
<td>0.60</td>
<td>4.82</td>
<td>20.50</td>
</tr>
<tr>
<td>LB + glycerol (4 g/L)</td>
<td>0.33</td>
<td>-</td>
<td>0.33</td>
<td>0.063</td>
<td>1.09</td>
<td>0.21</td>
<td>0.43</td>
<td>5.60</td>
<td>18.50</td>
</tr>
<tr>
<td>LB + glycerol (4 g/L) + glucose (4 g/L)</td>
<td>0.28</td>
<td>-</td>
<td>0.19</td>
<td>0.059</td>
<td>1.16</td>
<td>0.21</td>
<td>0.35</td>
<td>2.95</td>
<td>15.00</td>
</tr>
</tbody>
</table>

9.3.2. Model structure

In this study, a dynamic mathematical model was developed in batch cultures to capture the system characteristics and to design a feeding strategy. Following the development of the model in batch cultures, parameter estimation and model validation were performed. The optimal glucose feeding, induction time and initial optical density were predicted by the model based dynamic optimisation in fed-batch cultures to minimize the acetate production. At higher glucose feeding rates, acetate levels become inhibitory and at lower levels glucose starvation occurs. Monod kinetics was used for specific growth and death rate and also specific glucose consumption and acetate production rate in all identified functional states. The Monod model was developed by Nobel Laureate F. Monod in 1942 and for more than 60 years has been one of the most frequently used models in microbiology.

\[
\mu = \mu_{\text{max}} \left( \frac{S}{K_S + S} \right) \tag{9.1}
\]
$\mu$  Specific growth rate [h$^{-1}$]

$\mu_{\text{max}}$ Maximum specific growth rate [h$^{-1}$]

$S$  Concentration of the limiting substrate for growth [g/L]

$K_S$  Half-velocity constant- the value of $S$ when $\mu/\mu_{\text{max}} = 0.5$ [g/L]

The following assumptions are made for the models development of the considered fed-batch fermentations of *E. coli*:

- The bioreactor/flask is completely mixed.
- Potential mixing effects of the highly concentrated feeds with the fermentation medium are neglected for the sake of the model simplicity.
- The suspension viscosity in the reactor remains constant during the experiment.
- The substrate (glucose) is consumed mainly oxidatively.
- Variations in the growth rate, as well as in substrate consumption do not significantly change the elemental composition of biomass, thus balanced growth conditions are only assumed.
- Parameters, e.g. pH and temperature, are controlled to certain acceptable constant values during the process.
The structure of the model was developed based on following kinetics:

- Total mass balance:

\[
\frac{dV}{dt} = F_{in} - F_{out}
\]  \[9.2\]

- \( F_{in} \)  Flow rate of input [L/h]
- \( F_{out} \)  Flow rate of output [L/h]
- \( V \)  Broth volume [L]
- \( t \)  Time [h]
Mass balances on viable and dead cells:

\[
\frac{dVX_v}{dt} = (\mu - \mu_d)VX - F_{out}X_v \tag{9.3}
\]

\[
\frac{dVX_d}{dt} = \mu_dVX - F_{out}X_d \tag{9.4}
\]

- \(F_{out}\): Flow rate of output [L/h]
- \(V\): Broth volume [L]
- \(t\): Time [h]
- \(X\): Biomass concentration [g biomass/L]
- \(X_d\): Death biomass concentration [g biomass/L]
- \(X_v\): Viable biomass concentration [g biomass/L]
- \(\mu\): Specific growth rate [h\(^{-1}\)]
- \(\mu_d\): Specific Death rate [h\(^{-1}\)]

Growth rate:

\[
\mu = \mu_{max} \left( \frac{[GLC]}{K_{glc} + [GLC]} \right) \left( \frac{K_{i,ACO}}{K_{l,ACO} + [ACO]} \right) \tag{9.5}
\]

- \([GLC]\): Glucose concentration [g/L]
- \([ACO]\): Acetate concentration [g/L]
- \(K_{glc}\): Glucose coefficient for biomass production [g/L]
- \(K_{i,ACO}\): Acetate inhibition constant for growth [g/L]
- \(\mu\): Specific growth rate [h\(^{-1}\)]
- \(\mu_{max}\): Maximum specific growth rate [h\(^{-1}\)]
• Death rate:

\[ \mu_d = \mu_{d,max} \left( \frac{[AcO]}{K_{d,AcO} + [AcO]} \right) \]  \hspace{1cm} [9.6]

\([ACO]\) Acetate concentration [g/L]
\(K_{d,AcO}\) Acetate inhibition constant for death [g/L]
\(\mu_d\) Specific death rate [h\(^{-1}\)]
\(\mu_{d,max}\) Maximum specific death rate [h\(^{-1}\)]

• Glucose balance

\[
\frac{d(V[GLC])}{dt} = -V\mu_X Q_{x,glc} + F_{in}[GLC]_{in} - F_{out}[GLC] \]  \hspace{1cm} [9.7]

\[Q_{x,glc} = \left( \frac{\mu}{Y_{x,glc}} + M_{x,glc} \right) \]  \hspace{1cm} [9.8]

\([GLC]\) Glucose concentration [g/L]
\([GLC]_{in}\) Glucose concentration in input [g/L]
\(F_{in}\) Flow rate of input [L/h]
\(F_{out}\) Flow rate of output [L/h]
\(\mu\) Specific growth rate [h\(^{-1}\)]
\(X_v\) Viable biomass concentration [g biomass/L]
\(Q_{x,glc}\) Glucose consumption rate [h\(^{-1}\)]
\(Y_{x,glc}\) Yield of biomass from glucose [g biomass/g glucose]
\(M_{x,glc}\) Maintenance coefficient [g biomass/g glucose/h]
\(V\) Broth volume [L]
\(t\) Time [h]
In this study, a relatively simple mathematical model was developed for a typical industrial fermentation with the aim of producing high yield of cellulose, which incorporates acetate production and inhibition and includes the effect of glucose feeding. The sources for the parameters estimated for the development of the model are presented in Table 9.3. The values for maximum specific growth rate, glucose coefficient for biomass production and maintenance coefficient were taken from the literature [287, 288]. Maximum specific growth rate after induction was estimated based on data obtained from cultivation of GM C41 (DE3) cells induced by 0.05 mM of IPTG at 30 °C, which is presented in Figure 8.3, in previous chapter. Maximum specific death rate was calculated based on data presented in Figure 9.3, as the data captures 24-hr fermentation and exhibits a slight decrease after 12 hours in growth of the cultures supplemented with 10 g/L glucose and 4 g/L glycerol. The death rate of the cultures was calculated in a range of 0.012-0.0005 h⁻¹ and the average value was calculated approximately as 0.005 h⁻¹. The acetate inhibition constant for growth in fed-batch cultures was also obtained from literature as 25 g/L. Since the acetate inhibition constant for death was not
available in literature, a rough estimation was made for initial value and 50 g/L was used for the model the simulations. The yield of biomass from carbon source was estimated based on the lowest yield (0.21 g/g) obtained from experimental data presented in Table 9.2. Similarly, the lowest acetate yield obtained from the this data was used as yield of acetate from glucose

The values of the model parameters were re-estimated from batch data with a 95% confidence interval. All model and parameter estimation calculations were implemented in the advanced process modelling environment gPROMS® Model Builder 3.7.1. Model is still in the development and structural control stage since global sensitivity analysis has not been performed (Figure 9.7).

The model developed in this study is a dynamic and high-fidelity model, which is based on first principles starting directly from established laws of physics without making assumptions such as empirical or fitted parameters. It was established based on 9 equations (9.2-9.10), which contains 13 variables in total \( \{ V,F_{in},F_{out}, X, X_v, X_d, \text{[GLC]}, \text{[GLC]}_{in}, \text{[AcO]}, \mu, \mu_d, Q_{x,\text{glc}}, \text{and } Q_{\text{AcO,glc}} \} \). Three of the variables were set (\( \text{[GLC]}_{in}, F_{in} \) and \( F_{out} \)).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_{\text{max}} )</td>
<td>Maximum specific growth rate</td>
<td>0.7 [1/h]</td>
<td>[288]</td>
</tr>
<tr>
<td>( \mu_{\text{max}} )</td>
<td>Maximum specific growth rate (after induction)</td>
<td>0.23 [1/h]</td>
<td>Figure 8.3</td>
</tr>
<tr>
<td>( K_{glc} )</td>
<td>Glucose coefficient for biomass production</td>
<td>0.12 [mM]</td>
<td>[288]</td>
</tr>
<tr>
<td>( K_{i,\text{AcO}} )</td>
<td>Acetate inhibition constant for growth</td>
<td>25 [g/L]</td>
<td>[287]</td>
</tr>
<tr>
<td>( \mu_{d,\text{max}} )</td>
<td>Maximum specific death rate</td>
<td>0.005 [1/h]</td>
<td>Figure 9.3</td>
</tr>
<tr>
<td>( K_{d,\text{AcO}} )</td>
<td>Acetate inhibition constant for death</td>
<td>50 [g/L]</td>
<td>This study</td>
</tr>
<tr>
<td>( Y_{x,\text{glc}} )</td>
<td>Yield of biomass from glucose</td>
<td>0.21 [g/g]</td>
<td>Table 9.2</td>
</tr>
<tr>
<td>( M_{x,\text{glc}} )</td>
<td>Maintenance coefficient</td>
<td>0.03 [g/g/h]</td>
<td>[287]</td>
</tr>
<tr>
<td>( Y_{\text{AcO,glc}} )</td>
<td>Yield of acetate from glucose</td>
<td>0.35 [g/g]</td>
<td>Table 9.2</td>
</tr>
</tbody>
</table>
9.3.3. Model simulations

For each model simulations the model was given the following information: (i) initial broth volume; (ii) initial biomass, acetate and glucose concentrations; (iii) glucose feeding profile (if necessary).

For Experiment 1, initial conditions were presented in Table 9.4. The cells were cultured at 30°C in 200 ml flasks at 180 rpm, in the presence of 4.81 g/L glucose. The initial optical density of the culture was 0.05 at 600 nm. Glucose concentration was slightly higher due to experimental error compared to previous experiments reported in this chapter.

Table 9.4: Initial conditions for Experiment 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Broth Volume</td>
<td>200 ml</td>
</tr>
<tr>
<td>Initial Biomass Concentration</td>
<td>OD&lt;sub&gt;600&lt;/sub&gt; = 0.05</td>
</tr>
<tr>
<td>Glucose Concentration</td>
<td>4.81 [g/L]</td>
</tr>
<tr>
<td>Acetate concentration</td>
<td>0.08 [g/L]</td>
</tr>
</tbody>
</table>

Figure 9.8 and Figure 9.9 shows the comparison between model predictions and experimental data for Experiment 1. The arrow indicates the induction time in Figure 9.6. The model accurately predicted the growth profile and the concentrations of glucose and acetate. The growth curve and the glucose consumption exhibited a similar profile with previous experiment reported in this chapter, which was conducted in the presence of 4 g/L glucose. Final acetate concentration was detected slightly lower compared to this experiment, and it was still lower than the range of acetate concentrations that inhibits the growth completely (10-20 g/L).
Figure 9.8: Comparison of growth profile between model simulation and experimental data in batch cultures.

Figure 9.9: Comparison of glucose and acetate concentrations between model simulations and experimental data for Experiment 1.
In order to generate a high density culture, this predictive model was used to estimate the induction time for cultures with a higher initial concentration of cells. For this purpose, Experiment 2 conducted by inoculating the culture with a higher initial concentration of the cells.

The initial conditions for Experiment 2 were presented in Table 9.5. The cells were cultured at 30°C in 200 ml flasks at 180 rpm, in the presence of 4.86 g/L glucose. The initial OD$_{600}$ of the culture was increased to 0.12.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Broth Volume</td>
<td>200 ml</td>
</tr>
<tr>
<td>Initial Biomass Concentration</td>
<td>OD$_{600}$ = 0.12</td>
</tr>
<tr>
<td>Glucose Concentration</td>
<td>4.86 [g/L]</td>
</tr>
</tbody>
</table>

The comparison between model predictions and experimental data for Experiment 2 is presented in Figure 9.10. When the initial amount of the cell was increased, glucose was completely consumed and OD$_{600}$ reached 4.3 in 10 hours. The model was able to predict the induction time and the nutrient depletion time, accurately. Consequently, it was utilised to obtain the design of the feeding strategy for fed-batch cultures.
Fed-batch fermentation processes frequently allow significant increases in cell density for the achievement of higher product concentrations for recombinant systems. A fed batch experiment was designed based on the model developed to test this hypothesis in *E. coli* C41 (DE3) (pBCS + pCMP). The initial conditions for this experiment were the same with the initial conditions of Experiment 2, presented in Table 9.5. The cells were cultured at 30°C in 200 ml flasks at 180 rpm. Glucose feeding was started at 8 hours, which was followed by a second feeding at 10 hours.

9.3.4. Fed-Batch Cultures

*Figure 9.10:* Comparison of growth profile and glucose concentration between model simulations and experimental data for Experiment 2.
Figure 9.11: Comparison of growth profile and glucose concentration between model simulations and experimental data in fed-batch cultivation in flasks.

The model precisely captured the induction time, which was 3.5 hours after fermentation started, as presented in Figure 9.11. Glucose consumption and growth profile were also predicted accurately for the first 12 hours of the culturing. However, 3.2 g/L glucose was still abundant in the end of fermentation, in contrast to the model simulation, which predicted that glucose was exhausted in 14 hours. Since glucose was consumed very slowly in the cultures, feeding was performed two times. Also, final OD_{600} was obtained 3.5 experimentally; however, it was predicted as 10.6 by simulations. The glucose consumption was lower than predicted as a result of the inhibition in growth. The main reason behind this was attributed to the rapid decrease in DO levels. Oxygen availability influences cell growth negatively, which results in a significant decrease in the growth rate [273, 274]. Lack of DO control in flask experiments was the main reason for the poor cell growth. However model structure was built for experiments to be conducted in
bioreactor, assuming the existence of DO control. For this reason, the model was not effective in capturing late stage of growth profile and carbon source consumption.

Also, pH values lower than 5.5 significantly decreases the cell growth and even causes cell death (Table 9.1)[271]. Despite the manual control of pH of the culture medium by addition of NaOH (2M) to keep the pH constant higher than 6.0, it showed a rapid decrease during cultivation, as presented in Figure 9.12.

In an effort to achieve sufficient and accurate DO and pH control, a fed-batch experiment was performed in bioreactor.

![Figure 9.12: pH profile of fed-batch cultures.](image)
### Table 9.6: Initial conditions for fed-batch cultivation in the bioreactor.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>30°C</td>
</tr>
<tr>
<td>Broth Volume</td>
<td>4 L</td>
</tr>
<tr>
<td>Initial Biomass Concentration</td>
<td>(OD_{600} = 0.99)</td>
</tr>
<tr>
<td>Glucose Concentration</td>
<td>10.51 [g/L]</td>
</tr>
<tr>
<td>pH</td>
<td>7.0±0.25</td>
</tr>
</tbody>
</table>

The bioreactor was run at 30 °C and pH set at 7.0±0.25, as presented in Table 9.6. Dissolved oxygen (DO) was calibrated at 0% and set at 100%, obtained using 250 rpm agitation. After calibration, DO remained at approximately 100% until inoculation. Induction was performed at \(OD_{600} = 1.0\) by 0.05 mM IPTG and the cells allowed to grow overnight under pH and oxygen control. The media was supplemented with 10 g/L glucose in order to ensure that the glucose was not going to become exhausted overnight based on the model simulations (Figure 9.13).

The model was precisely able to estimate the induction and glucose exhaustion times; 4 hours and 12 hours, respectively. Glucose feeding was performed at 12 hours, which was followed by a second feeding at 13 hours. However, the growth rate was slower compared to the simulations predicted by the model, which resulted in a slower glucose consumption rate. Insufficient DO supply was obtained as the main technical problem in the bioreactor system utilised in this work, also mentioned in previous chapters as the main reason for unsatisfactory cell growth in bioreactor. The DO concentration in fed-batch cultivation of the cells conducted in bioreactor also fell to values that are below 2%. As a result, model predictions were not capturing the data obtained from the second half of the fermentation. Additionally, as the decrease in pH was extremely rapid correlated with sharp decrease in \(O_2\) levels, 455 ml of NaOH (4M) was used in 20 hours to maintain the pH at 7.0±0.25, which resulted in an approximately %12.5 dilution in culture media.
Although the cell growth was lower than expected, the cellulose concentration was higher, 40.0 mg/L, compared to fed-batch cultivation conducted in flasks, as presented in Table 9.7. Also, final cellulose concentration was higher in fed-batch cultures compared to concentrations obtained in batch cultures. However, the volumetric productivities in both conditions were very low compared to BC productivity in \textit{A. xylinum} cultures.

![Figure 9.13: Comparison of growth profile and glucose concentration between model simulations and experimental data in fed-batch cultivation in bioreactor.](image)

**Table 9.7: Comparison of growth characteristics of C41 (DE3) cells in fed-batch cultures.**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Specific growth rate (1/h)</th>
<th>Final DCW (g/L)</th>
<th>Yield of biomass from carbon source (g/g)</th>
<th>Yield of cellulose from carbon source (g/g)</th>
<th>Final cellulose concentration (mg/L)</th>
<th>Volumetric productivity (mg/L.h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flask</td>
<td>0.48</td>
<td>1.57</td>
<td>0.14</td>
<td>2.24</td>
<td>26.00</td>
<td>1.1</td>
</tr>
<tr>
<td>Bioreactor</td>
<td>0.30</td>
<td>1.82</td>
<td>0.16</td>
<td>3.64</td>
<td>40.00</td>
<td>1.7</td>
</tr>
</tbody>
</table>
9.3.5. Microscope Images and Direct Observations of Product

The morphology of the product was observed by light microscope at an appropriate magnification in this chapter. The product displayed an irregular pattern including cell aggregations around fibres, as presented in Figure 9.14. This is attributed to aggregative curli fimbriae formation as mentioned in Chapter 8, which occurs with BC production in *E. coli*. The structure of the substance exhibited a biofilm-like material exhibits rigid however fragile connections between cells [219]. Significant structures of the BC fibers were observed in extracellular material, which were an incredibly smooth and extremely long with a length of up to 2 cm, which are extremely longer compared to BC fibres produced by *A. xylinum*.

![Microscope Images](image)

*Figure 9.14*: Microscope images at different magnifications of the biofilm produced in this study (A-N). Microscope images of biofilm produced by *Salmonella* Typhimurium MAE52 (O and P), AFM images of biofilm produced by *Salmonella* Typhimurium MAE52 (R and S)[289].
Figure 9.14: Microscope images at different magnifications of the biofilm produced in this study (A-N). Microscope images of biofilm produced by *Salmonella* Typhimurium MAE52 (O and P). AFM images of biofilm produced by *Salmonella* Typhimurium MAE52 (R and S)[289].
Curli fimbriae, cellulose and cell aggregations are components in *Salmonella* biofilms as well as pathogenic and commensal *Escherichia coli* strains. The microscope images of the material produced in this study exhibited a similar structure the biofilm produced by *Salmonella Typhimurium MAE52* [289]. Image A shows how *E. coli* cells started to form aggregates, similary to the initial cell aggredeation in *Salmonella Typhimurium MAE52* in Image O. Following that, a
thicker and denser biofilm formation was detected in the material, as presented in Image B. This dense structure can be also seen in the biofilm produced by *Salmonella* Typhimurium MAE52 (Images P-S).

The extracellular substance exhibited a irregular surface structure, which contains embedded fibers inside and between the cell aggregations (Images C-J). The fibers displayed a very smooth structure with the length of up to 2 cm. Also, a thinner component with more curly and fragile connections was detected, as presented in Images K-L. The morphology of this substance is similar the biofilm produced by *E. coli* O157:H7 reported in previous chapter. (Figure 8.10 (5,6))[245, 246].

*Figure 9.15:* Direct observations of the biofilm produced in this study: length (a-b), liquid culture (c)
The extracellular substance secreted directly into culture media by bacteria, as also reported in previous chapters. After 5 hours culturing, the product was clearly noticeable by naked eye in the fermentation cultures of C41 (DE3) (pBCS + pCMP) as showed in Figure 9.15 (c). The substance was collected from culture by filtration and its size were obtained approximately up to 2 cm in length and 2 mm in width.

9.4. Discussion

Investigation of the growth limits of *E. coli* allows to convert it to a production machine for desired product. Bacteria would continue to grow and produce as long as their nutritional requirements are provided and no toxic or inhibitory factors are excreted into the medium.

In this chapter, the factors that have influence on bacterial growth was investigated including type and the concentration of carbon source, acetate inhibition, pH, oxygen and growth techniques (batch or fed-batch).

Effect of various concentrations of glucose and glycerol on growth, acetate accumulation and pH were investigated in batch cultures without pH control. Acetate accumulation occurs when the carbon flux, at the pyruvate junction of the central carbohydrate metabolism pathway, exceeds the capacity of the Krebs cycle (the Crabtree effect). Although cells exhibited slightly better growth profiles in the media provided by 10 g/L glucose, the acetate production was the highest with a concentration of 2.3 g/L. Replacement of glucose with glycerol resulted in lower acetate concentration, 1.6 g/L in the cultures supplemented with 4 g/L glycerol, which was in an agreement with literature [262, 263]. It has been demonstrected that *E. coli* BL21, also a derivative of *E. coli* B like *E. coli* C41 used in this study, produces very low amount of acetate in the presence of 2 g/L glucose concentration (0.35 g/L acetate), while *E. coli* JM109, a K12 derivative, produces 5 g/L acetate under the same conditions [118, 286]. This was attributed to the existence of an acetate control mechanism that operates in *E. coli* BL21 when
acetate concentration reaches beyond 1 g/L was, which directs the acetyl-CoA into glyoxylate cycle, which is a variation of TCA cycle to convert acetyl-CoA into succinic acid. Although this study also demonstrated that concentration of acetate accumulated in media was low, when the total amount of the cells in this study (OD$_{600}$=3.0-4.5) and in literature (OD$_{600}$=75.0-100.0) compared, acetate production per cell is significantly high. This could be the main reason for the inhibition of cell growth even there is still abundant carbon source in media. Another hypothesis is that the consumption of acetyl-CoA by glyoxylate cycle, which results in the production of succinic acid[290]. Also O$_2$ limitation initiates the formation of many metabolites of the acid metabolism such as succinate, fatty acids, acetaldehyde [276, 277]. This possibility is also in agreement with the rapid decrease in pH observed in cultures.

Although a limited amount of product was produced in this study, the highest concentration was obtained in the cultures supplemented with 4 g/L glucose in batch cultivation. Also, acetate production profiles were similar to each other in the presence of different concentrations of glucose and glycerol. Moreover, supplementing culture media with glucose higher than 5 g/L is inhibits the basal expression before induction to reduce the metabolic stress on the cells. For this reason, glucose used in the study reported in this chapter and glucose levels were maintained lower than 5 g/L after the induction.

A dynamic model developed for the feeding strategy of glucose in effort to achieve a high density culture for a higher yield of product [79, 291]. Many papers have been published on improving the production of recombinant proteins from microorganisms; however, most of them do not capture such factors as acetate inhibition and cell death. The model developed in this study is a relatively simple model with 9 parameters incorporating the acetate inhibition and cell death, which allowed predicting glucose consumption, acetate production and induction time for batch cultures. On the other hand, the experiments conducted in flasks with glucose feeding and pH control did not follow the model predictions especially for the second half of the fermentation period.
The cell growth was inhibited possibly as a result of oxygen limitation, which also resulted in rapid decrease in pH with the accumulation of toxic metabolites. When the cells were cultured in stirred-tank bioreactor with feeding, insufficient DO supply was the major problem again. It is apparent that low DO concentrations, as well as oxygen fluctuations could be detrimental for protein expression [292]. Therefore, any significant improvement in BC concentrations was not observed.

The morphology of the product was detected by light microscopy in this chapter. The images collected in this study showed that the product of cellulose in E. coli is in a combination with cell aggradations. This was also proved in literature before [115, 116, 219, 221, 289]. However, in this study, fibres were very smooth and longer compared to biofilm-like products produced by E. coli.

### 9.5. Conclusion

The work reported in this chapter showed that BC in combination with cell aggradations was produced. The level of acetate and decrease in pH caused an inhibition in cell growth despite the carbon source feeding. As the system is a highly complex recombinant expression system, any small change in these parameters could strongly affect the productivity. On the other hand, high plasmid stability and cell growth in an expression system requires a satisfactory oxygen supplement. In contrast, biofilm like materials is produced under low agitation or static conditions. This could be another reason for the low concentrations of product. Further investigation for the optimisation of culture conditions should be done to achieve high density cultures for higher productions.
10. GENERAL CONCLUSIONS AND FUTURE WORK
10.1. General Conclusions

Bacterial cellulose has been gaining significant attention for a wide range of industrial application as a green material with outstanding chemical physical and mechanical properties. However, BC yields are not justifiable for the market because of the high costs as a consequence of the slow growth rate of *A. xylinum*, the most efficient BC producer. Therefore, it is essential to investigate new valuable strains for the improvement of BC production. However, the genetic pathway of BC biosynthesis requires many enzymatic steps and the functions of the components required for BC production such as CmcAx and CcpAx are not fully identified. Although the necessity of improvement of BC production in *E. coli* has been mentioned in the literature several times, complexity of the system was a big challenge for scientists. In this study, we have developed a novel recombinant expression system in *E. coli* for BC production based on the genes essential for the biosynthesis. Following three different *E. coli* strains have been utilized to investigate the BC production: BL21 (DE3), HMS174 (DE3) and C41 (DE3). The results showed that the production and the quality of the material are highly strain-specific. The main outcomes of the research described in the thesis are described, as follows.

Recombinant expression of *bcs* operon and its upstream operon was achieved by two plasmids: pBCS and pCMP. Both plasmid vectors are developed by using the backbone for Duet vectors. It has been proved in literature that Duet vectors allow co-expression of multiple genes successfully. However, the target operons are extremely long products for cloning, which risk the stability of DNA. In effort to overcome this, they were generated by artificial gene synthesis. This allowed complete sequence verification, optimization of codon use and GC content of the target sequences. The orientation of *bcs* operon in pBCS and the upstream operon in pCMP were verified by restriction digest test. Furthermore, the full sequence details of target sequences and cloning strategy were provided. Simultaneous transformation of pBCS and pCMP into *E. coli* cells was achieved by chemical transformation.
As a consequence of the optimisation of GC contents and codon usage according to *E. coli*, the translational efficiencies of genes of interest were detected in high levels by q-PCR and SDS-PAGE analysis in BL21 (DE3). Although genetically modified (GM) strain BL21 (DE3) is very efficient for the high throughput production of recombinant proteins, this caused an extremely high metabolic burden on bacteria, which resulted in a sharp decrease in the percentage of plasmid bearing cells. Fermentation of GM BL21 (DE3) was conducted in the presence of various IPTG concentrations (0.025, 0.05, 0.1, and 0.2), and at various temperatures (22, 30, and 37 °C). However, neither any significant improvement in plasmid stability nor BC production was detected. On the other hand, genetically modified (GM) strain HMS174 (DE3) exhibited very high plasmid stability above 85% after 10 hours fermentation at 22 °C or 30°C. Consequently, the product formation was detected after 24 hours fermentation using this novel expression system, at 22 °C in the presence of 0.025 mM IPTG. The OD$_{600}$ of GM BL21 (DE3) cultures reached approximately 3.0, at 22 °C in the presence of 0.025 mM IPTG, whereas this was only 2.2 in the cultures of GM HMS174 (DE3) at the end of 10 hours culturing. Limited growth of GM HMS174 (DE3) leaded the further investigations to focus on a superior strain in membrane protein expression: *E. coli* C41 (DE3). The percentage of plasmid bearing cells was observed extremely high in this system, which was above 90% even after 24 hours. Cell aggradation and product formation were detected 5 hours after IPTG induction. These findings showed that a highly stable expression system was established for BC production in *E. coli*.

A fed-batch system was developed based on a dynamic model incorporating the acetate inhibition and cell death, which successfully achieved to predict glucose consumption, acetate production and induction time. Feeding strategy was design based on the model, however sharp decreases in O$_2$ and high acetate accumulations were observed in the cultures, which negatively influenced the cell growth and volumetric yields.
Maximum specific growth rates, final concentrations of dry cell weights (dcw) and product concentrations are presented in Table 10.1. Among GM cellulose producing strains, best performing cell line was obtained as GM C41 (DE3). The highest growth rate was detected as 0.48 h\(^{-1}\) in fed batch cultivation of GM C41 (DE3) conducted in flaks when the cells were induced by 0.05 mM IPTG at 30 °C. The highest product concentration and final dcw were obtained as 1.82 g/L and 40.0 mg/L, respectively, when this cell line was cultured in fed-batch bioreactor under the same conditions. The volumetric productivity was calculated as 1.7 mg/L.h, when the highest cellulose concentration was achieved in fed-batch fermentation of GM C41 (DE3) in bioreactor. Volumetric productivity of \textit{A. xylinum} ATCC 53582, which was used as the donor organism for \textit{bcs} genes in this study, was reported between 5.95 -16 mg/L.h [69, 293]. The volumetric productivity achieved in this study approximately 10 times lower than the original producer of cellulose due to poor O\(_2\) supply. However, fermentation period in this study is only 24 hours, whereas this was reported as 7 days for \textit{A. xylinum} cultures. Therefore, the GM strain has a great potential to produce cellulose with less energy consumption.

The morphology of BC produced by HMS174 (DE3) and C41 (DE3) were characterised by SEM. BC produced by HMS174 (DE3) exhibited an incredible fibre structure. BC fibrils displayed distinct and smooth features and created a dense network, which shows a similar structure with BC produced by \textit{A. xylinum}. The fibres were approximately 400 times longer and 200 times wider than BC cellulose produced by \textit{A. xylinum}. BC produced by C41 (DE3) exhibited a more random and biofilm-like structure, which also displayed cell-cell interactions. The fibres with the length of approximately 600 μm were detected, which were embedded into the cell-cell interactions. On the other hand, microscope images revealed the significant structure of the BC fibers in extracellular material, which exhibited a very smooth structure with the length of up to 2 cm.
Table 10.1: Summary of maximum specific growth rates, dcw and product concentrations under various conditions studied in this work.

<table>
<thead>
<tr>
<th>Temperature</th>
<th>IPTG Concentration</th>
<th>37°C</th>
<th>30°C</th>
<th>22°C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>μmax (1/h)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL21 (DE3)</td>
<td>un-induced</td>
<td>0.47</td>
<td>0.42</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>0.025 mM</td>
<td>0.42</td>
<td>0.42</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>0.05 mM</td>
<td>0.42</td>
<td>0.38</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>0.1 mM</td>
<td>1.20</td>
<td>1.31</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>final dcw (g/l)</td>
<td>0.96</td>
<td>0.62</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>product concentration (mg/l)</td>
<td>-</td>
<td>-</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>0.025 mM (bioreactor)</td>
<td>0.43</td>
<td>0.40</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.2 mM</td>
<td>1.20</td>
<td>1.31</td>
<td>2.59</td>
</tr>
<tr>
<td>HMS174 (DE3)</td>
<td>un-induced</td>
<td>0.45</td>
<td>0.36</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>0.025 mM</td>
<td>0.36</td>
<td>0.33</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>0.2 mM</td>
<td>1.15</td>
<td>0.57</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>final dcw (g/l)</td>
<td>1.15</td>
<td>0.57</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>product concentration (mg/l)</td>
<td>-</td>
<td>-</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>0.025 mM (bioreactor)</td>
<td>0.43</td>
<td>0.40</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>0.2 mM</td>
<td>1.20</td>
<td>1.31</td>
<td>2.59</td>
</tr>
<tr>
<td>C41 (DE3)</td>
<td>un-induced</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>0.025 mM</td>
<td>0.32</td>
<td>0.30</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>0.2 mM</td>
<td>0.99</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>final dcw (g/l)</td>
<td>0.99</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>product concentration (mg/l)</td>
<td>-</td>
<td>-</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>0.025 mM (bioreactor)</td>
<td>0.32</td>
<td>0.30</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Although the concentrations of BC were observed low and need to be elucidated further, hence, the novel biosystems developed in this study could provide stable and scalable for valuable products with incredible fibre structure for further investigations on BC and the study could be potentially used in the improvement of the BC production for industrial purposes with shorter cultivation periods to reduce energy consumption. Furthermore, biofilms are also as a robust new platform, which allows the design of nanomaterials for the purpose of waste water treatment, manufacture of pharmaceuticals, fabricate new textiles, etc.
10.2. Future work

This study presented a novel study on the development of a recombinant expression system for the production of bcs proteins and the establishment of a stable bioprocess for the BC production. Although this is an important step from the scientific point of view, further investigations should be continued on improvement of the production for industrial use. Future work should involve enhancement of BC production by optimisation of bioprocess parameters including culturing techniques, nutrients, dissolved oxygen, pH and agitation of culture medium. Furthermore, identification of the genetic pathway of cellulose biosynthesis is also highly important to have a better understanding of the regulation of the biosynthesis.

The use of glucose feeding with an amino-acid containing feed stream has proven a very effective strategy for both preventing the acetate formation and enhancing the growth rate incredibly [294]. Similarly, supplementing the media with glycine and methionine or glycine could reduce the acetate formation and improve the protein expression [295]. Furthermore, when fructose was used as a carbon source in replacement of glucose, a reduction in acetate formation and a 65% increase in protein production were observed [296]. In fed-batch cultures, the lowering of the feed rate of the carbon source could restrict formation of acetate [118, 297]. It has been also reported that supplement of vitamins and trace elements to the media enhance the growth rate [298]. An initiation step could enhance the initiation of polymerisation or crystallization of a polymer. It has been proved that supplementing the media with chitosan enhanced the attachment of BC to the chitosan to generate a composite material [299]. Therefore, addition of cellulose to the media could improve the BC production and could also provide a surface for genetically modified E. coli C41 (DE3) to attach and stimulate the adhesion of biofilms, BC, and cell aggregations to this surface. On the other hand, biofilm formation, which occurs in combination with BC production in E. coli, is observed as a response for stress conditions such as temperature, osmolarity, pH and oxygenation [300]. However, stress conditions could inhibit the cell growth.
dramatically. As high agitation results in cellulose non-producing mutants among *A. xylinum* populations, use of low agitation speeds could be a potential solution to stimulate the cell aggradation [301]. Also, high agitated cultures result in highly branched, three dimensional, reticulated structure, whereas low agitated cultures produce a normal cellulose pellicle with a lamellar structure and less significant branching. However, low agitation causes a decrease in oxygen supply for medium. Too meet all the demands, these factors should be carefully optimised and balanced to maintain cell growth and improve the production simultaneously.

Since the morphologies of BC samples produced by genetically modified strains HMS174 (DE3) and C41 (DE3) displayed rather different structures, use of different strains compatible with pBCS and pCMP should be also studied for the further investigation of potential cellulose-producing strains. It has been proved that cellulose production occurs only when *adrA*, a gene encoding a putative transmembrane protein, is expressed in *E. coli* [116]. Moreover, diguanyl cyclase (DGC) protein, which synthesizes c-di-GMP (DGC), play a key role for the activation of cellulose-biosynthesis in bacteria [272]. Both proteins encoded by short sequences, which could be cloned into pCMP to stimulate the biosynthesis of BC without causing any significant metabolic burden on genetically modified cells. In addition, cellulose biosynthesis mechanism in *E. coli* requires further investigations as it has not been fully understood.

High density cultures are major requirement for enhancing the overall yield of product. The constant supply of fresh nutrients and removal of metabolic waste such as acetate could favour an enhanced growth of all genetically modified strains HMS174 (DE3) and C41 (DE3), hence increasing the concentration of biosynthesised BC. Development of mathematical models for the tight control of the feeding and induction mechanism is highly effective for the achievement of high density cultures. However in this study, model predictions were not effective for capturing the data obtained from the late stages of fermentation since it was based on the assumption of efficient O2 supply throughout fermentation of the cells. In order to solve this problem, the DO supply system of the bioreactor utilised in this work
should be replaced with an efficient DO supply system. This solution potentially could have a significant effect on cell growth and product concentration. Also, the model should be modified by addition of equations describing the effect of DO dynamics on cell growth, in order to control the system fully.
REFERENCES


atcc23769 and atcc53582: Comparison of cellulose synthetic ability between strains. DNA Research, 2002. 9(5): p. 149-156.


110. A. Abbott, Bacterial cellulose for use in: Hierarchical composites, macroporous foams, bioinorganic nanohybrids and bacterial-based nanocomposites. 2011, Imperial College London.


160. K.A. Zingaro and E.T. Papoutsakis, Toward a semisynthetic stress response system to engineer microbial solvent tolerance. MBio, 2012. 3(5).


Appendix 1

BSA calibration curve

\[ y = 0.0012x \]

\[ R^2 = 0.9867 \]

\( \text{OD}_{595} \)

\( \mu g/ml \)
Appendix 2

Correlation factor of dry weights to optical densities

\[ y = 2.2826x \]

\[ R^2 = 0.9976 \]
Appendix 3

a. Full sequence of pCDFDuet-1:

1. GGGGAATTGT GACGGGATAA CAATTCCCCT GTAGAAATAA TTTTGTTTAA CTTTAATAAG
2. GAGATATACC ATGGGCAGCA GCCATCACCA TCATCACCAC AGCCAGGATC CGAATTCGAG
3. CTCGGCGCGC CTGCAGGTCG ACAAGCTTGC GGCCGCATAA TGCTTAAGTC GAACAGAAAG
4. TAATCGTATT GTACACGGCC GCATAATCGA AATTAATACG ACTCACTATA GGGGAATTGT
5. GAGCGGATAA CAATTCCCCA TCTTAGTATA TTAGTTAAGT ATAAGAAGGA GATATACATA
6. TGGCAGATCT CAATTGGATA TCGGCCGGCC ACGCGATCGC TGACGTCGGT ACCCTCGAGT
7. CTGGTAAAGA AACCGCTGCT GCGAAATTTG AACGCCAGCA CATGGACTCG TCTACTAGCG
8. CAGCTTAATT AACCTAGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA CCCCTTGGGG
9. CCTCTAAACG GGTCTTGAGG GGTTTTTTGC TGAAACCTCA GGCATTTGAG AAGCACACGG
10. TCACACTGCT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
11. TCACACTGCT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
12. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
13. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
14. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
15. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
16. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
17. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
18. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
19. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
20. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
21. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
22. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
23. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
24. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
25. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
26. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
27. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
28. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
29. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
30. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
31. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
32. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
33. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
34. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
35. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
36. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
37. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
38. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
39. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
40. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
41. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
42. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
43. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
44. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
45. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
46. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
47. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
48. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
49. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
50. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
51. TCAACACTGT CTCGTTATGC AATAACCGGG TAAACCCGCA AATAGCTATG CGTGAGCAGG
b. Full sequence of pETDuet-1:

1  GGGGAATTGT  GACGCCGTTAA  CAAATTTCCCCT  CGTAGAAATAA  TTTTGTTTAA  CTTTAAGAAG
61  GAGATATACC  ATGGGCAGCA  GCCATCACCA  TCATCACCAC  AGCCAGGATC  CGAATTCGAG
121  CTCGGCGCGC  CTGCAGGTCG  ACAAGCTTGC  GGCCGCATAA  TGCTTAAGTC  GAACAGAAAG
181  TAATCGTATT  GTACACGGCC  GCATAATCGA  AATTAATACG  ACTCACTATA  GGGGAATTGT
241  GAGCGGATAA  CAATTCCCCA  TCTTAGTATA  TTAGTTAAGT  ATAAGAAGGA  GATATACATA
301  TGGCAGATCT  CAATTGGATA  TCGGCCGGCC  ACGCGATCGC  TGACGTCGGT  ACCCTCGAGT
361  CTGGTAAAGA  AACCGCTGCT  GCGAAATTTG  AACGCCAGCA  CATGGACTCG  TCTACTAGCG
421  CAGCTTAATT  AACCTAGGCT  GCTGCCACCG  CTGAGCAATA  ACTAGCATAA  CCCCTTGGGG
481  CCTCTAAACG  GGTCTTGAGG  GGTTTTTTGC  TGAAAGGAGG  AACTATATCC  GGATTGGCGA
541  ATGGGACGCG  CCCTGTAGCG  GCGCATTAAG  CGCGGCGGGT  GTGGTGGTTA  CGCGCAGCGT
601  GACCGCTACA  CTTGCCAGCG  CCCTAGCGCC  CGCTCCTTTC  GCTTTCTTCC  CTTCCTTTCT
661  CGCCACGTTC  GCCGGCTTTC  CCCGTCAAGC  TCTAAATCGG  GGGCTCCCTT  TAGGGTTCCG
721  ATTTAGTGCT  TTACGGCACC  TCGACCCCAA  AAAACTTGAT  TAGGGTGATG  GTTCACGTAG
781  TGGGCCATCG  CCCTGATAGA  CGGTTTTTCG  CCCTTTGACG  TTGGAGTCCA  CGTTCTTTAA
841  TAGTGGACTC  TTGTTCCAAA  CTGGAAACAAC  ACTCAACCCT  ATCTCGGTCT  ATTCTTTTGA
901  TTTATAAGGG  ATTTTGCCGA  TTTCGGCCTA  TTGGTAAAAA  AATGAGCTGA  TTTAACAAAA
961  ATTTAACGCG  AATTTTAACA  AAATATTTAAC  GTTTACAATT  TCTGGCGGCA  CGATGGCATG
1021  AGATTATCAA  AAAGGATCTT  CACCTAGATC  CTTTTAAATT  AAAAATGAAG  TTTTAAATCA
1081  ATCTAAAGTA  TATATGGATG  AACTTGGTAC  TACGTTTAC  AATGCTTAAT  CAGTGAGGCA
1141  CCTATCTCAG  CGATCTGTCT  ATTTCGTTCA  TCCATAGTTG  CCTGACTCCC  CGTCGTGTAG
1201  ATAACTACGA  TACGGGAGGG  CTTACCATCT  GGCCCCAGTG  CTGCAATGAT  ACCGCGAGAC
1261  CCACGCTCAC  CGGCTCCAGA  TTTATCAGCA  ATAAACCAGC  CAGCCGGAAG  GGCCGAGCGC
1321  AGAAGTGGTC  CTGCAACTGC  ATCCAGCCTC  TATCCAGCCTA  TAACTTGGT  CGCGAAGCT
1381  AGAATAGTGA  GTGCGCCAGT  TAATAGTTTG  CGCGATTTTC  TGCATCTGG  TACCGACAT
1441  GTGCGGTCAC  CGTCTGGCTT  TGGTATGGCT  TCATTCAGCT  CCGGTTCCCA  ACGATCAAGG
1501  CGAGTTACAT  GATCCCCCAT  GTTGTGCAAA  AAAGCGGTTA  GCTCCTTCGG  TCCTCCGATC
1561  GTGTGTCAGAA  GTAAGTTGGC  CGCAGTGTTA  TCACTCATGG  TTATGGCAGC  ACTGCATAAT
1621  TCTCTACCTG  TACGTCACAT  GTGAAGATCG  TTTTCGTGAA  CTGTGAGGTA  CTCAACCAAG
1681  TCTATTCTAG  AATAGTGTAG  TTCCTTCGTT  GTGCTCATCA  TTGGAAAACG  TTCTTCGGGG
1741  CGAAAACTCT  CAAGGCTGTTG  ACCGCTGTTG  AGATCCAGTT  CGATGTAACC  CACTCGTGCA
1801  CCCAACTGAT  CTTCAGGCTA  CTTTCTTAAA  AATGCTTAAT  CAGTGAGGCA
1861  AGGCAAAATG  CGGACGTAAA  GGGGTAATTG  GGGCGGCAAG  CTGCCAGTAT  ACTCTGCTCA
1921  CCTCTGTTGCT  TGGTATGGCT  TCATTCAGCT  CCGGTTCCCA  ACGATCAAGG
1981  AGAAGTGGTC  CTGCAACTGC  ATCCAGCCTC  TATCCAGCCTA  TAACTTGGT  CGCGAAGCT
2041  AGAATAGTGA  GTGCGCCAGT  TAATAGTTTG  CGCGATTTTC  TGCATCTGG  TACCGACAT
2101  GTGCGGTCAC  CGTCTGGCTT  TGGTATGGCT  TCATTCAGCT  CCGGTTCCCA  ACGATCAAGG
2161  TAGACATCGC  AAGGAGGATC  CCGACAGTCG  GTAACCGGCT  TCTCAGTTGC  CTTTGTTTCG
2221  CGAGTTCAGA  CTGAGGGTAA  CAGCCCTGCT  GTCCACGTCTC  TCTCGTGGCT  CTTTGTTTCG
2281  AGATACCAAA  TATTGACATG  AACTTGGTAC  TACGTTTAC  AATGCTTAAT  CAGTGAGGCA
2341  TAGACATCGC  AAGGAGGATC  CCGACAGTCG  GTAACCGGCT  TCTCAGTTGC  CTTTGTTTCG
2401  ATTAAGATGC  ATGGGCAACT  GTGCTGTGCTT  GCCGCTGCTT  AACCAAGCAA  ACCTCACTGC
2461  CTCCAGCTTG  CGTCCAGGAT  TACGTCGCTT  CTGCTGTGCTT  GCCGCTGCTT  AACCAAGCAA
2521  CGAAGAATCTT  CACGATGTC  TCAACGTTGGT  GATCCATGAC  CTGTCATGAC  CTGTCATGAC
2581  CCCAGCTGAT  CTTTATGATG  GAGGCGAGCTT  AAGGCGAGCTT  CAGTCGCTT  AAGGCGAGCTT
2641  CCACAATGTC  CGAGATCGTC  TCTCAGGCTT  CTGCTGTGCTT  GCCGCTGCTT  AACCAAGCAA
2701  TTTGTTGATG  CTTGTCAGTG  GGGCAGGAGC  TGGTAAAGATG  CGAGATCGTC  TCTCAGGCTT
2761  TACGTCGCTT  CGTCCAGGAT  TACGTCGCTT  CTGCTGTGCTT  GCCGCTGCTT  AACCAAGCAA
2821  ATTCATGTCG  TAAGATCCAT  CAGGGCTGAT  AAGGAGGATC  CCGACAGTCG  GTAACCGGCT
2881  CGAGTTCAGA  CTGAGGGTAA  CAGCCCTGCT  GTCCACGTCTC  TCTCGTGGCT  CTTTGTTTCG
2941  GTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3001  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3061  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3121  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3181  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3241  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
3301  CGTCTGACCGA  GCAGGTTCGA  ATTTCATGCC  GAAGATCGTC  TCTCAGTTGC  CTTTGTTTCG
### Appendix 4

**a. Zero-Cutter enzymes for bcsAB:**

<table>
<thead>
<tr>
<th>#</th>
<th>Enzyme</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AatII</td>
<td>G__ACGT_T</td>
</tr>
<tr>
<td>2</td>
<td>AgeI</td>
<td>A__CCGG_AT</td>
</tr>
<tr>
<td>3</td>
<td>Ascl</td>
<td>GG__CCGG_AC</td>
</tr>
<tr>
<td>4</td>
<td>AseI</td>
<td>AT__TA__AT</td>
</tr>
<tr>
<td>5</td>
<td>AsaI</td>
<td>GCG__AT__CGC</td>
</tr>
<tr>
<td>6</td>
<td>AvrII</td>
<td>C__CTAG__G</td>
</tr>
<tr>
<td>7</td>
<td>Bael</td>
<td>(N)_5 (N)_10ACNNNNGTAYC(N)_7 (N)_5</td>
</tr>
<tr>
<td>8</td>
<td>BbvCI</td>
<td>CC__TCA__GC</td>
</tr>
<tr>
<td>9</td>
<td>BglII</td>
<td>A__GATC_T</td>
</tr>
<tr>
<td>10</td>
<td>BmrI</td>
<td>ACTGGGNNNN_N</td>
</tr>
<tr>
<td>11</td>
<td>BmtI</td>
<td>G__CTAG__C</td>
</tr>
<tr>
<td>12</td>
<td>BsuEI</td>
<td>CTTGAG(N)_5 (N)_10NN</td>
</tr>
<tr>
<td>13</td>
<td>BsaXI</td>
<td>NNN_N_NAC(N)_2 CTCC(N)_7 NNN</td>
</tr>
<tr>
<td>14</td>
<td>BsiWI</td>
<td>C__GTAC__G</td>
</tr>
<tr>
<td>15</td>
<td>BsmBl</td>
<td>CGTCTCN_NNNN</td>
</tr>
<tr>
<td>16</td>
<td>BspHI</td>
<td>T__CATG_A</td>
</tr>
<tr>
<td>17</td>
<td>BspQI</td>
<td>GCTCTTCN_NNNN</td>
</tr>
<tr>
<td>18</td>
<td>BssHII</td>
<td>G__CGCG__C</td>
</tr>
<tr>
<td>19</td>
<td>Dral</td>
<td>TTT__AAA</td>
</tr>
<tr>
<td>20</td>
<td>Drdi</td>
<td>GACNN_NN_NNGTC</td>
</tr>
<tr>
<td>21</td>
<td>Eagl</td>
<td>C__GGCC__G</td>
</tr>
<tr>
<td>22</td>
<td>EarI</td>
<td>CTCTTCN_NNNN</td>
</tr>
<tr>
<td>23</td>
<td>EcI</td>
<td>GCCGGA(N)_5 (N)_10NN</td>
</tr>
<tr>
<td>24</td>
<td>Eco53KI</td>
<td>GAG__CTC</td>
</tr>
<tr>
<td>25</td>
<td>EcoRI</td>
<td>G__AATT__C</td>
</tr>
<tr>
<td>26</td>
<td>FseI</td>
<td>GG__CCGG__CC</td>
</tr>
<tr>
<td>27</td>
<td>Hpal</td>
<td>GTT__AAC</td>
</tr>
<tr>
<td>28</td>
<td>MfeI</td>
<td>C__AATT__G</td>
</tr>
<tr>
<td>29</td>
<td>MluI</td>
<td>A__CCGG__AT</td>
</tr>
<tr>
<td>30</td>
<td>NdeI</td>
<td>CA__TA__TG</td>
</tr>
<tr>
<td>31</td>
<td>Nhel</td>
<td>G__CTAG__C</td>
</tr>
<tr>
<td>32</td>
<td>NotI</td>
<td>GC__GGCC__GC</td>
</tr>
<tr>
<td>33</td>
<td>PaeRI</td>
<td>TTA__AT__TAA</td>
</tr>
<tr>
<td>34</td>
<td>PaeR7I</td>
<td>C__TCGA__G</td>
</tr>
<tr>
<td>35</td>
<td>PflFI</td>
<td>GACN__N_NGTC</td>
</tr>
<tr>
<td>36</td>
<td>Pmel</td>
<td>GTTT__AAC</td>
</tr>
<tr>
<td>37</td>
<td>PsiI</td>
<td>TTA__TAA</td>
</tr>
<tr>
<td>38</td>
<td>PspXI</td>
<td>VC__TCGA__GB</td>
</tr>
</tbody>
</table>
### Zero-Cutter enzymes for bcsCD:

<table>
<thead>
<tr>
<th>#</th>
<th>Enzyme</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AatII</td>
<td>GAACGT C</td>
</tr>
<tr>
<td>2</td>
<td>AccI</td>
<td>GT MKAC</td>
</tr>
<tr>
<td>3</td>
<td>AflII</td>
<td>CT TAAG</td>
</tr>
<tr>
<td>4</td>
<td>AhdI</td>
<td>GACNN N NNGTC</td>
</tr>
<tr>
<td>5</td>
<td>ApaI</td>
<td>GGGGCC C</td>
</tr>
<tr>
<td>6</td>
<td>AscI</td>
<td>GG CCGGGCC</td>
</tr>
<tr>
<td>7</td>
<td>AsiSI</td>
<td>GCGATCGC</td>
</tr>
<tr>
<td>8</td>
<td>AvrII</td>
<td>C CTAGG</td>
</tr>
<tr>
<td>9</td>
<td>BamHI</td>
<td>G GATCC C</td>
</tr>
<tr>
<td>10</td>
<td>BanII</td>
<td>G RGCCY C</td>
</tr>
<tr>
<td>11</td>
<td>BbsI</td>
<td>GAAAGACNN NNNN</td>
</tr>
<tr>
<td>12</td>
<td>BbvCl</td>
<td>CC TCA GGC</td>
</tr>
<tr>
<td>13</td>
<td>BcoVI</td>
<td>GTATCC(N)6 N</td>
</tr>
<tr>
<td>14</td>
<td>BglII</td>
<td>A GATCC T</td>
</tr>
<tr>
<td>15</td>
<td>BmlI</td>
<td>GCTAG C</td>
</tr>
<tr>
<td>16</td>
<td>BpuEl</td>
<td>CTTGAG(N)14 NN</td>
</tr>
<tr>
<td>17</td>
<td>BsaBI</td>
<td>GATNN NNATC</td>
</tr>
<tr>
<td>18</td>
<td>Bsal</td>
<td>GGTCTCN NNNN</td>
</tr>
<tr>
<td>19</td>
<td>BseRI</td>
<td>GAGGAG(N)22 NN</td>
</tr>
<tr>
<td>20</td>
<td>BsiWI</td>
<td>C GTAGC</td>
</tr>
<tr>
<td>21</td>
<td>BsmBI</td>
<td>CGTCTCN NNNN</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Recognition Site</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td></td>
</tr>
<tr>
<td>BspDI</td>
<td>ATCGAT</td>
<td></td>
</tr>
<tr>
<td>BspH1</td>
<td>TCCATG</td>
<td></td>
</tr>
<tr>
<td>BspQI</td>
<td>GCCTCTCN NNN</td>
<td></td>
</tr>
<tr>
<td>BssHII</td>
<td>GC GGCG C</td>
<td></td>
</tr>
<tr>
<td>BstBI</td>
<td>TT CG AA</td>
<td></td>
</tr>
<tr>
<td>BstZ17I</td>
<td>GTA TAC</td>
<td></td>
</tr>
<tr>
<td>ClaI</td>
<td>AT CGAT</td>
<td></td>
</tr>
<tr>
<td>Dral</td>
<td>TTTAAA</td>
<td></td>
</tr>
<tr>
<td>Drdl</td>
<td>GACNNNN NNGTC</td>
<td></td>
</tr>
<tr>
<td>EagI</td>
<td>CGGCCG</td>
<td></td>
</tr>
<tr>
<td>Eco53kl</td>
<td>GAG CTC</td>
<td></td>
</tr>
<tr>
<td>EcoRI</td>
<td>GAAATT C</td>
<td></td>
</tr>
<tr>
<td>Mmel</td>
<td>TCCRAC(N)NN NN</td>
<td></td>
</tr>
<tr>
<td>Ncol</td>
<td>CATG G</td>
<td></td>
</tr>
<tr>
<td>Ndel</td>
<td>CA TAG TG</td>
<td></td>
</tr>
<tr>
<td>Nhel</td>
<td>GCTAG C</td>
<td></td>
</tr>
<tr>
<td>NotI</td>
<td>GCCGGGC GC</td>
<td></td>
</tr>
<tr>
<td>NruI</td>
<td>TCG CGA</td>
<td></td>
</tr>
<tr>
<td>NsiI</td>
<td>AAGTGCA T</td>
<td></td>
</tr>
<tr>
<td>PacI</td>
<td>TTA AT TAA</td>
<td></td>
</tr>
<tr>
<td>PfFI</td>
<td>GACNN NGTGC</td>
<td></td>
</tr>
<tr>
<td>Pmel</td>
<td>GTTT AAAC</td>
<td></td>
</tr>
<tr>
<td>PsII</td>
<td>TTA TAA</td>
<td></td>
</tr>
<tr>
<td>PsPOMI</td>
<td>GGGCC C</td>
<td></td>
</tr>
<tr>
<td>PspXI</td>
<td>VC TCGA GB</td>
<td></td>
</tr>
<tr>
<td>Rsrl</td>
<td>CG GWC CG</td>
<td></td>
</tr>
<tr>
<td>SacI</td>
<td>GAGCT C</td>
<td></td>
</tr>
<tr>
<td>SacII</td>
<td>CGGC GG</td>
<td></td>
</tr>
<tr>
<td>Sall</td>
<td>GT CGA C</td>
<td></td>
</tr>
<tr>
<td>SapI</td>
<td>GCCTCTCN NNN</td>
<td></td>
</tr>
<tr>
<td>SfiI</td>
<td>GGCCCN NNN NGGCC</td>
<td></td>
</tr>
<tr>
<td>SnaBl</td>
<td>TAC GTA</td>
<td></td>
</tr>
<tr>
<td>SpeI</td>
<td>A CTAG T</td>
<td></td>
</tr>
<tr>
<td>SphI</td>
<td>GATCG C</td>
<td></td>
</tr>
<tr>
<td>SrlI</td>
<td>GCCGGG GCC</td>
<td></td>
</tr>
<tr>
<td>SspI</td>
<td>AAT ATT</td>
<td></td>
</tr>
<tr>
<td>Swal</td>
<td>ATTT AAAAA</td>
<td></td>
</tr>
<tr>
<td>TfiI</td>
<td>G AWT C</td>
<td></td>
</tr>
<tr>
<td>Tth111I</td>
<td>GACN NN NGTC</td>
<td></td>
</tr>
<tr>
<td>Xbal</td>
<td>TCTAG A</td>
<td></td>
</tr>
<tr>
<td>XmnI</td>
<td>GAANN NNTTC</td>
<td></td>
</tr>
<tr>
<td>ZraI</td>
<td>GAC GTC</td>
<td></td>
</tr>
</tbody>
</table>
### a. Zero-Cutter enzymes for ccpAx:

<table>
<thead>
<tr>
<th>#</th>
<th>Enzyme</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AatII</td>
<td>G↓ACGT C</td>
</tr>
<tr>
<td>2</td>
<td>Acc65I</td>
<td>G GTAC↓C</td>
</tr>
<tr>
<td>3</td>
<td>Accl</td>
<td>GT MKAC</td>
</tr>
<tr>
<td>4</td>
<td>Acul</td>
<td>CTGAAG(N)₁₀NN</td>
</tr>
<tr>
<td>5</td>
<td>Afel</td>
<td>AGC↓GCT</td>
</tr>
<tr>
<td>6</td>
<td>AffII</td>
<td>C TTAA↓G</td>
</tr>
<tr>
<td>7</td>
<td>Ahdl</td>
<td>GACNN↓N NNGTC</td>
</tr>
<tr>
<td>8</td>
<td>Apal</td>
<td>G↓GGCC C</td>
</tr>
<tr>
<td>9</td>
<td>ApaLI</td>
<td>G TGCA↓C</td>
</tr>
<tr>
<td>10</td>
<td>Apol</td>
<td>R AATT↓Y</td>
</tr>
<tr>
<td>11</td>
<td>Ascl</td>
<td>GG CGCG↓CC</td>
</tr>
<tr>
<td>12</td>
<td>AseI</td>
<td>AT TA↓AT</td>
</tr>
<tr>
<td>13</td>
<td>AseSI</td>
<td>GCG↓AT CGC</td>
</tr>
<tr>
<td>14</td>
<td>AvrII</td>
<td>C CTAG↓G</td>
</tr>
<tr>
<td>15</td>
<td>Bael</td>
<td>(N)₁₀ACNNGTAYC(N)₇(N)₅</td>
</tr>
<tr>
<td>16</td>
<td>BamHI</td>
<td>G GATC↓C</td>
</tr>
<tr>
<td>17</td>
<td>BbsI</td>
<td>GAAGACNN NNNN</td>
</tr>
<tr>
<td>18</td>
<td>BbvCI</td>
<td>CC TCA↓GC</td>
</tr>
<tr>
<td>19</td>
<td>Bcgl</td>
<td>↓NN (N)₁₀CGA(N)₆TGC(N)₁₀NN</td>
</tr>
<tr>
<td>20</td>
<td>BclVI</td>
<td>GTATCC↓C</td>
</tr>
<tr>
<td>21</td>
<td>BclI</td>
<td>T GATC↓A</td>
</tr>
<tr>
<td>22</td>
<td>BcoDI</td>
<td>GTCTCN NNNN</td>
</tr>
<tr>
<td>23</td>
<td>BglII</td>
<td>GCCN↓NN NNGC</td>
</tr>
<tr>
<td>24</td>
<td>Blpl</td>
<td>GC TNA↓GC</td>
</tr>
<tr>
<td>25</td>
<td>Bmrl</td>
<td>ACTGGGNNN↓N</td>
</tr>
<tr>
<td>26</td>
<td>BmtI</td>
<td>G↓CTAG C</td>
</tr>
<tr>
<td>27</td>
<td>BpmI</td>
<td>CTGGAG(N)₁₄NN</td>
</tr>
<tr>
<td>28</td>
<td>Bpu10I</td>
<td>CC TNA↓GC</td>
</tr>
<tr>
<td>29</td>
<td>BpuEl</td>
<td>CTTGAG(N)₁₄NN</td>
</tr>
<tr>
<td>30</td>
<td>BsaAI</td>
<td>YAG↓GTR</td>
</tr>
<tr>
<td>31</td>
<td>BsaBI</td>
<td>GATNN↓NNATC</td>
</tr>
<tr>
<td>32</td>
<td>BsaHI</td>
<td>GR CG↓YC</td>
</tr>
<tr>
<td>33</td>
<td>Bsal</td>
<td>GGTCTCN NNNN</td>
</tr>
<tr>
<td>34</td>
<td>BseRI</td>
<td>GAGGAG(N)₁₀NN</td>
</tr>
<tr>
<td>35</td>
<td>BseYI</td>
<td>C↓CCAG↓C</td>
</tr>
<tr>
<td>36</td>
<td>BsiWI</td>
<td>C GTAC↓G</td>
</tr>
<tr>
<td>37</td>
<td>BsmAI</td>
<td>GTCTCN NNNN</td>
</tr>
<tr>
<td>38</td>
<td>BsmBI</td>
<td>CGTCTCN NNNN</td>
</tr>
<tr>
<td>39</td>
<td>BsmFI</td>
<td>GGGAC(N)₁₀ NNNN</td>
</tr>
<tr>
<td>40</td>
<td>BspDI</td>
<td>AT CG↓AT</td>
</tr>
<tr>
<td></td>
<td>Enzyme</td>
<td>Sequence</td>
</tr>
<tr>
<td>---</td>
<td>--------</td>
<td>----------------</td>
</tr>
<tr>
<td>41</td>
<td>BsrBI</td>
<td>CCGCTC</td>
</tr>
<tr>
<td>42</td>
<td>BsrDI</td>
<td>GCAATGNN</td>
</tr>
<tr>
<td>43</td>
<td>BsrGI</td>
<td>TGTACA</td>
</tr>
<tr>
<td>44</td>
<td>BstHII</td>
<td>CGCGC</td>
</tr>
<tr>
<td>45</td>
<td>BstBI</td>
<td>TTCCAA</td>
</tr>
<tr>
<td>46</td>
<td>BstXI</td>
<td>CCANNNNNTGG</td>
</tr>
<tr>
<td>47</td>
<td>BstZ17I</td>
<td>GTAC</td>
</tr>
<tr>
<td>48</td>
<td>BsuCI</td>
<td>GGATGNN</td>
</tr>
<tr>
<td>49</td>
<td>ClaI</td>
<td>ATCGAT</td>
</tr>
<tr>
<td>50</td>
<td>CspCl</td>
<td>(NN)(N)CAA(N)GTGG(N)NN</td>
</tr>
<tr>
<td>51</td>
<td>CviQI</td>
<td>GTC</td>
</tr>
<tr>
<td>52</td>
<td>Dral</td>
<td>TTTAAA</td>
</tr>
<tr>
<td>53</td>
<td>DralII</td>
<td>CACNNNNTG</td>
</tr>
<tr>
<td>54</td>
<td>DrdI</td>
<td>GACNNNNNGTC</td>
</tr>
<tr>
<td>55</td>
<td>EagI</td>
<td>CGGCCG</td>
</tr>
<tr>
<td>56</td>
<td>Eco53kl</td>
<td>GAGCTC</td>
</tr>
<tr>
<td>57</td>
<td>EcoNI</td>
<td>CCTNNNNNNGAGG</td>
</tr>
<tr>
<td>58</td>
<td>EcoO109I</td>
<td>RGNCICY</td>
</tr>
<tr>
<td>59</td>
<td>EcoRI</td>
<td>GAACTC</td>
</tr>
<tr>
<td>60</td>
<td>EcoRV</td>
<td>GATATC</td>
</tr>
<tr>
<td>61</td>
<td>FokI</td>
<td>GGATGNN</td>
</tr>
<tr>
<td>62</td>
<td>FseI</td>
<td>GGCCGGCC</td>
</tr>
<tr>
<td>63</td>
<td>HindIII</td>
<td>AGCTT</td>
</tr>
<tr>
<td>64</td>
<td>Kasi</td>
<td>GCCGC</td>
</tr>
<tr>
<td>65</td>
<td>KpnI</td>
<td>GTAC</td>
</tr>
<tr>
<td>66</td>
<td>MfII</td>
<td>CATTG</td>
</tr>
<tr>
<td>67</td>
<td>NaeI</td>
<td>GGCCGC</td>
</tr>
<tr>
<td>68</td>
<td>Ndel</td>
<td>CATA</td>
</tr>
<tr>
<td>69</td>
<td>NheI</td>
<td>GTAGC</td>
</tr>
<tr>
<td>70</td>
<td>NotI</td>
<td>GCGGCCGC</td>
</tr>
<tr>
<td>71</td>
<td>NsiI</td>
<td>ATGCA</td>
</tr>
<tr>
<td>72</td>
<td>Pali</td>
<td>TTAAT</td>
</tr>
<tr>
<td>73</td>
<td>PciI</td>
<td>CATG</td>
</tr>
<tr>
<td>74</td>
<td>PfI</td>
<td>GCNNNGTC</td>
</tr>
<tr>
<td>75</td>
<td>PfMII</td>
<td>CCANNNTGG</td>
</tr>
<tr>
<td>76</td>
<td>PluII</td>
<td>GCCGC</td>
</tr>
<tr>
<td>77</td>
<td>Pmel</td>
<td>GTTTAAAC</td>
</tr>
<tr>
<td>78</td>
<td>PmlII</td>
<td>CACGTG</td>
</tr>
<tr>
<td>79</td>
<td>PpuMI</td>
<td>RGWGCACY</td>
</tr>
<tr>
<td>80</td>
<td>Psil</td>
<td>TTAAT</td>
</tr>
<tr>
<td>81</td>
<td>PspOMI</td>
<td>GGGCCGC</td>
</tr>
<tr>
<td>82</td>
<td>PspXI</td>
<td>VCTGCAGGB</td>
</tr>
<tr>
<td>83</td>
<td>PvuII</td>
<td>CAGCTG</td>
</tr>
<tr>
<td>84</td>
<td>Rsal</td>
<td>GTAC</td>
</tr>
<tr>
<td></td>
<td>Enzyme</td>
<td>Sequence</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>-------------------</td>
</tr>
<tr>
<td>85</td>
<td>Rsrl</td>
<td>CG GWC CG</td>
</tr>
<tr>
<td>86</td>
<td>SacI</td>
<td>G AGCT C</td>
</tr>
<tr>
<td>87</td>
<td>SalI</td>
<td>G TCGA C</td>
</tr>
<tr>
<td>88</td>
<td>SbfI</td>
<td>CC TGCA GG</td>
</tr>
<tr>
<td>89</td>
<td>Scal</td>
<td>AGT ACT</td>
</tr>
<tr>
<td>90</td>
<td>SexAI</td>
<td>A CCWGG GT</td>
</tr>
<tr>
<td>91</td>
<td>Sfil</td>
<td>GGCCN NNN NGGCC</td>
</tr>
<tr>
<td>92</td>
<td>SfoI</td>
<td>GCC GGC</td>
</tr>
<tr>
<td>93</td>
<td>SnaBl</td>
<td>TAC GTA</td>
</tr>
<tr>
<td>94</td>
<td>Spel</td>
<td>A CTAG GT</td>
</tr>
<tr>
<td>95</td>
<td>SrfI</td>
<td>GCCC GGGC</td>
</tr>
<tr>
<td>96</td>
<td>SspI</td>
<td>AAT ATT</td>
</tr>
<tr>
<td>97</td>
<td>Stul</td>
<td>AGG CCT</td>
</tr>
<tr>
<td>98</td>
<td>Swal</td>
<td>ATTT AAAT</td>
</tr>
<tr>
<td>99</td>
<td>Tfil</td>
<td>G AWT C</td>
</tr>
<tr>
<td>100</td>
<td>Tth111I</td>
<td>GACN N NGTC</td>
</tr>
<tr>
<td>101</td>
<td>Xbal</td>
<td>T CTAG A</td>
</tr>
<tr>
<td>102</td>
<td>XcmI</td>
<td>CCANNN N NNNNTGG</td>
</tr>
<tr>
<td>103</td>
<td>Zral</td>
<td>GAC GTC</td>
</tr>
</tbody>
</table>
Appendix 5

a. Full sequence and GC% content of cmcax:

1  GGATCCGATG  AGTGTGATGG  CAGCCATGGG  CGGTGCACAG  GTGCTGAGCA  GCACAGGCGC
61 TGCGCGCGCG  AGCGCGTCG  TGGCTGTGAC  CCAGCAAGGT  GCGCAAAGGC  ATAGCGAAGG
121 GCCGGCTGAC  GGGCGACAC  ACCTCGACC  AAGAAGAACT  CAGGACAGCA  ATGCCGCATT
181 CTAGTCATG  GAGCGGCGCG  ATCGCGCTCC  GCAACGGGAC  TTAAAGAATG  TGGCGAGTG
241 CCACGCGCGG  TGTGGTACTG  CTGAGCCAGG  CCAAGGAGAG  ACCCGGGCGG  TCAGCCCGCG
301 GCAGCGCGCC  CGCGGGTCAT  GAGGCACAGC  TGGCCTGCTG  GGTGAGGACT  TCAGTTAGC
361 GTGGCATATTG  TGGGTTGGGG  TGAGTGCTGA  AGAGGGGAGG  AGCCAGGTTG  CTGGTGTGTA
421 AGCGCGGTAT  CCGGCGCTCC  GCGCGGGTGT  AAGGCCACGT  TGGATCGACT  CTGGGTCTGA
481 GACAGCCTTTG  AAGCGGCGCG  ATCGGCGCTA  TGGTGTGTGC  GGTGAGGACT  TCAGTTAGC
541 TCGCAGCGTC  GCGGAGGCG  ACGCGCCCCA  GAGGCGCGCG  GCGGAGGCG  CCGGAGCAGC
601 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
661 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
721 TGGGTTGGGG  TGAGTGCTGA  AGAGGGGAGG  AGCCAGGTTG  CTGGTGTGTA  GACAGCCTTT
781 ACGGCGGTAT  CCGGCGCTCC  GCGCGGGTGT  AAGGCCACGT  TGGATCGACT  CTGGGTCTGA
841 AGCGCGGTAT  CCGGCGCTCC  GCGCGGGTGT  AAGGCCACGT  TGGATCGACT  CTGGGTCTGA
901 TCGCAGCGTC  GCGGAGGCG  ACGCGCCCCA  GAGGCGCGCG  GCGGAGGCG  CCGGAGCAGC
961 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
1021 GGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG

b. Full sequence and GC% content of ccpAx:

1  AGATCTCATG  AGCGCGCCAG  GTAGTGACGA  AGTGGCCCGT  GGTGGCCAGG  CAGGTAGTCC
61 TCAGGACTTC  CAGGCGCGCT  TACGCGCTCT  GCAGGGACG  GGCGCGCAGT  AGACGTACCG
121 TCCGGTGTGAG  GACGCGCGCT  TGACGCGCTG  CGCGCGCGCG  GACCGGCGCG  AGGTCGCTTT
181 CGCAGCGGCGG  ACGGCAGCGG  CGCGCGCGCG  GACCGGCGCG  AGGTCGCTTT  GCAGCGCGCG
241 AGCAGCGCGG  TGTGGTACTG  CTGAGCCAGG  CCAAGGAGAG  ACCCGGGCGG  TCAGCCCGCG
301 GCAGCGCGCC  CGCGGGTCAT  GAGGCACAGC  TGGCCTGCTG  GGTGAGGACT  TCAGTTAGC
361 GTGGCATATTG  TGGGTTGGGG  TGAGTGCTGA  AGAGGGGAGG  AGCCAGGTTG  CTGGTGTGTA
421 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
481 GACAGCCTTTG  AAGCGGCGCG  ATCGGCGCTA  TGGTGTGTGC  GGTGAGGACT  TCAGTTAGC
541 TCAGGACTTC  CAGGCGCGCT  TACGCGCTCT  GCAGGGACG  GGCGCGCAGT  AGACGTACCG
601 TCAGGACTTC  CAGGCGCGCT  TACGCGCTCT  GCAGGGACG  GGCGCGCAGT  AGACGTACCG
661 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
721 TCTTCGCTTCG  AGCGCGCGCG  GCAAGCGGCGG  CAGATCGAGG  AAGGTTCCTCG  GGTGCGCGCG
781 ACGGCGGTAT  CCGGCGCTCC  GCGCGGGTGT  AAGGCCACGT  TGGATCGACT  CTGGGTCTGA
841 AGCGCGGTAT  CCGGCGCTCC  GCGCGGGTGT  AAGGCCACGT  TGGATCGACT  CTGGGTCTGA
901 TCGCAGCGTC  GCGGAGGCG  ACGCGCCCCA  GAGGCGCGCG  GCGGAGGCG  CCGGAGCAGC
961 CGCGCGCGCG  TACCCGCGCG  GCGTCTACAT  CTTGCCGGTG  AGCAGGCGGT  GCCGAGTGTG
1021 TCGGCTTGGAG  GCAATAGCTGC  GCCGCGCGCG  GAAATGAGAAA  GAGATTTAAGC  TCGG
c. Full sequence and GC% content of bcsAB:

```
1 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
61 GGGTAAGATC CTGAGCATCC GCGGCGCAGG TCTGACCATT GGCGTGTTCG GCTTATGCGC
121 CCTGATCGCC GCAACAAGCG TTACCCTGCC GCCTGAGCAG CAGTTAATTG TTGCATTTGT
181 CGATGGCGTG ACCTGGCGTG ACAGCGGTCA TCGATGATTC CAGAAGGCAT GCAGAAGACC
241 GAGCTTCGAT ACCTGGCTGC AGGGCCTGCT GGGTACAATG CTGCTGGTTG CAGAACTGTA
301 CGCCCGTGAT ATGGCTGTTCC TGAGCTACTT CCAAACCATC GCCCCGTTAC ATCGTGCCCC
361 GTTACCGTTA CCTCGGAATC CGGATGAGTG GCCTACCGTT GACATCTTCG TTCCGACCTA
421 CAATGAAGAA TTAAGTATCG TGCGTCTGAC AGTGCTGGGC AGCCTGGGTA TCGACTGGCC
481 TCCGGAGAAG GTGCGCGTGC ACATCCTGGA CGATGGTCGC CGTCCGGAAT TTGCCGCCTT
541 TGCCGCCGAA TGCGGCGCAA ATTACATCGC ACGTCCTACC AACGAACACG CCAAGGCCGG
601 CAACTTAAAC TACGCCATCG GCCACACCGA TGGCGACTAC ATCCTGATCT TCGACTGTGA
661 CCATGTGCCG ACCCGCGCAT TCTTACAACT GACCATGGGT TGGATGGTGG AGGACCCGAA
721 GATTGCCCTG ATGCAAACCC CGCACCATTT CTACAGTCCG GACCCGTTTC AGCGCAATCT
781 GAGCGCAGGC TACCGCACAC CTCCGGAAGG TAACCTGTTC TATGGCGTGG TGCAGGACGG
841 AGCCACCGAG CGTCTGATCT TACACATCGG TCAGCGTGTT CGTTGGGCCC GCGGTATGCT
901 GCAGATTTTC CGCATCGACA ATCCGCTGTT TGGCCGCGGT TTAAGTTGGG GCCAGCGCCT
961 GTGTTACCTG AGCGCCATGA CAAGTTTCCT GTTCGCAGTG CCGCGCGTGA TCTTTCTGAG
1021 CAGCCCGCTG GCCTTCCTGT TCTTCGGCCA GAACATCATC GCAGCCAGCC CGTTAGCCCT
1081 GAGGCGCTGC TACAGCTTCT GGAGTGAGGT GTACGAAACC ACAATGGCCC TGTTTCTGGT
1141 GCGTGTTACC ATCGTGACAC TGCTGAGTCC GAGCCGCGGC AAATTTAACG TTACAGATAA
1201 AGGCGGTCTG CTGGAGAAGG GCTATTTCGA CCTGGGCGCC GTGTATCCGA ACATCATCCT
1261 GGGCCTGATC ATGTTTGGCG GTCTGGCACG CGGCGTTTAC GAACTGAGCT TCGGCCATCT
1321 GGACCAGATT GCCGAACGCG CCTACCTGTT AAACAGCGCC TGGGCCATGC TGAGTCTGAT
1381 CACCGGCAAC ACCAACCGGA AGCCAGTGAC CTGCCGCCGC TGCCGGCACT GCTGCCTGCA
1441 ACAAGCGGTG CAGCACAAGC CGGCGCAGGT GATGCAGGTG CCAATGGCCC GGGCAGCCCT
1501 ACAGGTCAGC CGTTAGCAGC CGATAGCGCC GACGCATTAG TGGAGAATGC CGAGAACACC
1561 AGCGATACAG CCACAGTGCA CAACTATACC CTGAAGGACC TGGGTGCCGC CGGTAGCATC
1621 ACAATGCGTG GTTTAGCCCC TCTGCAAGGC ATCGAGTTCG GCATTCCGAG TGATCAGCTG
1681 GGACCAGATT GCCGAACGCG CCTACCTGTT AAACAGCGCC TGGGCCATGC TGAGTCTGAT
1741 TATTCCTGCC ACAATTCCGG TTGAGGTGGC AAACGCAGAC GGCAGCATCA TTGTTACAGG
1801 GTGTTTCTCC ACAAATTCGG TTAGGAGTGC TGGCGCGGGG AGGCGCGCGG GGGCAAAACT
1861 AGCTGCAGAG ATACGCGGCC ATACGCGGCC ATACGCGGCC ATACGCGGCC ATACGCGGCC
1921 GCGACCGATT ATCGCGGAGT GTATGCGGAG CTGGCGGAGT GTATGCGGAG CTGGCGGAGT
1981 CAGCACGATT TTATGCGGAG TTTTCGCGGA ATACGCGGCC ATACGCGGCC ATACGCGGCC
2041 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2101 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2161 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2221 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2281 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2341 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2401 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2461 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2521 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2581 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2641 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2701 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2761 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2821 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2881 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
2941 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
3001 GGATCCGATG CCGAAGTGC GCAGCAGCAC ACAGAGCGAG AGTGGCATGA GCCAGTGGAT
```
d. Full sequence and GC% content of bcsAB:

1
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
61
AACAATGGTG
GCCGCGCTGT
TATTTAAGGC
AAACGGTGCA
CGCGCCCAGC
AGGCAGCAGA
121
AGCACAGGCA
CCTGCAAGCA
GACCACCAT
GATGCAGGCA
GCCACCGTGG
CACCGGACA
181
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
241
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
301
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
361
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
421
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
481
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
541
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
601
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
661
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
721
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
781
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
841
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
901
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
961
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
1021
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
1081
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
1141
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
1201
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
1261
GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC

Base Position

GC%  

0.0  

0.2  

0.4  

0.6  

0.8  

1.0  

GGCCGGCCAC
ATGACACACA
AGCGCTACGC
AAGCAGTCTG
AGCGCCGGCT
TACTGGCAAC
294

1321 TGAAGCGAA GCAACATCC CGGCAATCCT GGGCAATCAG GGGCAACCT GGAGCGCATC
1381 GGACCGGCTA GGAACGTGTA GAAGCCATG GCGGCAATCC GGGCAACTG GGAGCGCATC
1441 GAGGCGGCTA GGAACGTGTA GACGCGGCTA GGGCAACCT GGAGCGCATC
1501 GCGAGAACC GCGACGCCG GCGCCGCTG TATTAATCAT CGGCAACCT GGAGCGCATC
1561 GCGAGAACC GCGACGCCG GCGCCGCTG TATTAATCAT CGGCAACCT GGAGCGCATC
1621 GCGAGAACC GCGACGCCG GCGCCGCTG TATTAATCAT CGGCAACCT GGAGCGCATC
1681 TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT
1741 CGTCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT
1801 CGTCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT
1861 TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT
1921 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
1981 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2041 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2101 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2161 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2221 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2281 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2341 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2401 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2461 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2521 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2581 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2641 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2701 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2761 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2821 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2881 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
2941 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3001 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3061 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3121 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3181 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3241 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3301 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3361 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3421 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3481 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3541 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3601 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3661 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3721 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3781 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3841 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3901 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
3961 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4021 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4081 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4141 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4201 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4261 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4321 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4381 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG
4441 AGACTGATTT TCGCCGATTG TACAGCGGAT TCGCCGATTG TACAGCGGAT TCGCCGATTG

GC Content

Base Position

G%
Appendix 6
a. Sequence alignments of original and optimised cmcax:
78.1% identity in 1042 residues overlap; Score: 4796.0; Gap frequency: 0.0%

cmcax_ori
cmcax_opt

1 GG ATCCGATGTCGGTCATGGCGGCGATGGGAGGGGCGCAGGTGCTTTCATCCACCGGTGC
1 GGATCCGATGAGTGTGATGGCAGCCATGGGCGGTGCACAGGTGCTGAGCAGCACAGGCGC
**********
** ***** ** ***** ** ** ********
*** ** **

cmcax_ori
cmcax_opt

61 GTTCGCAGACACCGCCCCCGATGCGGTCGCGCAGCAATGGGCCATCTTCCGCGCCAAGTA
61 CTTCGCAGACACCGCACCTGACGCAGTGGCACAGCAGTGGGCCATCTTCCGTGCCAAGTA
************** ** ** ** ** ** ***** ************** ********

cmcax_ori
cmcax_opt

121 TCTTCGTCCCAGCGGACGTGTCGTGGATACGGGCAATGGTGGCGAATCCCATAGTGAGGG
121 TCTGCGCCCGAGCGGTCGTGTGGTTGATACCGGCAACGGTGGCGAGAGCCATAGCGAAGG
*** ** ** ***** ***** ** ***** ***** ********
****** ** **

cmcax_ori
cmcax_opt

181 GCAGGGCTATGGCATGCTCTTTGCCGCGTCGGCGGGGGACCTTGCGTCGTTCCAGTCGAT
181 CCAGGGCTACGGCATGCTGTTCGCCGCAAGCGCAGGTGACCTGGCCAGCTTCCAGAGTAT
******** ******** ** *****
** ** ***** **
******
**

cmcax_ori
cmcax_opt

241 GTGGATGTGGGCGCGCACCAACCTGCAGCATACCAATGACAAGCTGTTTTCCTGGCGGTT
241 GTGGATGTGGGCCCGCACCAACCTGCAGCACACCAACGACAAGCTGTTCAGTTGGCGCTT
************ ***************** ***** ***********
***** **

cmcax_ori
cmcax_opt

301 CCTCAAGGGGCATCAGCCCCCGGTGCCCGACAAGAACAATGCCACAGATGGCGACCTGCT
301 CCTGAAGGGTCATCAGCCGCCGGTGCCGGACAAGAACAATGCCACCGACGGTGACCTGCT
*** ***** ******** ******** ***************** ** ** ********

cmcax_ori
cmcax_opt

361 GATCGCGCTTGCGCTTGGTCGTGCGGGCAAGCGTTTCCAGCGCCCCGATTACATTCAGGA
361 GATCGCCTTAGCCCTGGGCCGTGCAGGTAAGCGCTTCCAACGCCCGGACTATATCCAGGA
****** * ** ** ** ***** ** ***** ***** ***** ** ** ** *****

cmcax_ori
cmcax_opt

421 CGCCATGGCCATTTATGGCGATGTGCTGAACCTGATGACGATGAAGGCGGGACCGTATGT
421 TGCCATGGCCATCTACGGCGACGTGCTGAACCTGATGACCATGAAAGCCGGTCCGTACGT
*********** ** ***** ***************** ***** ** ** ***** **

cmcax_ori
cmcax_opt

481 CGTCCTCATGCCCGGTGCTGTCGGCTTTACCAAGAAGGACAGCGTGATCCTCAACCTGTC
481 GGTTCTGATGCCGGGCGCCGTGGGCTTCACCAAGAAGGACAGCGTGATCCTGAACTTAAG
** ** ***** ** ** ** ***** *********************** *** *

cmcax_ori
cmcax_opt

541 CTATTACGTCATGCCCTCGCTGCTGCAGGCGTTCGACCTTACGGCCGACCCGCGCTGGCG
541 CTACTACGTGATGCCGAGTCTGCTGCAGGCCTTTGACCTGACAGCCGATCCGCGTTGGCG
*** ***** *****
*********** ** ***** ** ***** ***** *****

cmcax_ori
cmcax_opt

601 TCAGGTGATGGAAGACGGGATTCGCCTTGTTTCCGCCGGCCGTTTCGGGCAGTGGCGCCT
601 TCAGGTGATGGAGGACGGTATCCGCCTGGTGAGTGCAGGCCGTTTCGGCCAATGGCGTCT
************ ***** ** ***** **
** *********** ** ***** **

cmcax_ori
cmcax_opt

661 GCCCCCCGACTGGCTGGCGGTGAATCGCGCCACCGGTGCGCTGTCGATCGCATCGGGATG
661 GCCTCCGGATTGGTTAGCCGTGAACCGTGCAACCGGCGCCTTAAGCATCGCAAGCGGCTG
*** ** ** *** * ** ***** ** ** ***** ** *
******
** **

cmcax_ori
cmcax_opt

721 GCCGCCGCGCTTTTCCTATGATGCGATTCGGGTGCCGCTTTATTTTTATTGGGCGCATAT
721 GCCGCCGCGCTTTAGCTACGATGCCATCCGCGTGCCGCTGTACTTCTACTGGGCACACAT
************* *** ***** ** ** ******** ** ** ** ***** ** **

cmcax_ori
cmcax_opt

781 GCTGGCGCCGAACGTGTTGGCTGATTTCACCCGATTCTGGAATAATTTCGGGGCTAATGC
781 GCTGGCACCGAACGTGCTGGCCGACTTTACCCGCTTCTGGAACAATTTCGGCGCAAATGC
****** ********* **** ** ** ***** ******** ******** ** *****

cmcax_ori
cmcax_opt

841 CCTGCCAGGATGGGTTGATCTGACAACAGGGGCGCGTTCGCCGTACAACGCCCCGCCTGG
841 ACTGCCTGGTTGGGTTGACCTGACCACAGGTGCCCGCAGCCCTTACAACGCCCCTCCGGG
***** ** ******** ***** ***** ** **
** *********** ** **

cmcax_ori
cmcax_opt

901 ATATCTTGCTGTTGCCGAATGCACGGGGCTTGATTCTGCCGGGGAACTCCCGACACTGGA
901 TTACCTGGCCGTGGCAGAATGCACCGGTTTAGATAGCGCCGGCGAATTACCGACCTTAGA
** ** ** ** ** ******** ** * ***
***** *** * ***** * **

295


b. Sequence alignments of original and optimised ccpAx:

74.9% identity in 1075 residues overlap; Score: 4883.0; Gap frequency: 0.0%

cmcax_ori   961 TCATCGCCGCCATTATATTCCGCAGCGTTGACGCTGCTCGTTTACATCGCGCGGGCGGA
cmcax_opt   961 CCACGCCCCGATTAATATTCCGCAGCGTTGACGCTGCTCGTTTACATCGCGTGCCGA

**** **** *** ** ** ***** ** ** ***** ** ** ** **

cmcax_ori   1021 GGAGACTATAAAGTGAAAGCTT
cmcax_opt   1021 GGAAACCATCAAGTAAAAGCTT

*** **** ** ****

b. Sequence alignments of original and optimised ccpAx:

74.9% identity in 1075 residues overlap; Score: 4883.0; Gap frequency: 0.0%

cmcax_ori   961 TCATCGCCGCCATTATATTCCGCAGCGTTGACGCTGCTCGTTTACATCGCGCGGGCGGA
cmcax_opt   961 CCACGCCCCGATTAATATTCCGCAGCGTTGACGCTGCTCGTTTACATCGCGTGCCGA

**** **** *** ** ** ***** ** ** ***** ** ** ** **

cmcax_ori   1021 GGAGACTATAAAGTGAAAGCTT
cmcax_opt   1021 GGAAACCATCAAGTAAAAGCTT

*** **** ** ****
c. Sequence alignments of original and optimised bcsA:

77.7% identity in 2172 residues overlap; Score: 9920.0; Gap frequency: 0.0%

bcsA_ori 1 ATGCCAGAGGTTCGGTCGTCAACGCAGTCAGAGTCAGGAATGTCACAGTGGATGGGGAAA
bcsA_opt 1 ATGCCGGAAGTGCGCAGCAGCACACAGAGCGAGAGTGGCATGAGCCCGCGCCCTAG

bcsA_ori 61 ATTCTTTCCATTCGCGGTGCTGGGCTGACTATTGGTGTTTTTGGCCTGTGTGCGCTGATT
bcsA_opt 61 ATCCTGAGCATCCGCGGCGCAGGTCTGACCATTGGCGTGTTCGGCTTATGCGCCCTGATC

bcsA_ori 121 GCGGCTAGTCCGTGACCCGTGCGCCAGAACAGCAGTTGATTGTGGCATTTGTATGTGTC
bcsA_opt 121 GCCGCAACAAGCGTTACCCTGCCGCCTGAGCAGCAGTTAATTGTTGCATTTGTGTGCGT

bcsA_ori 181 GTGATCTTTTTTATTGTCGGTCATAAGCCCAGCCGTCGGTCCCAGATTTTCCTTGAAGTG
bcsA_opt 181 GTGATCTTTTTCATCGTGGGTCACAAGCCGAGTCGTCGCAGCCAAATTTTCCTGAGGTTG

bcsA_ori 241 CTGTCAGGGCTGGTTTCGCTGCGCTATCTGACATGGCGCCTGACGGAAACGCTTTCATTC
bcsA_opt 241 CTGAGCGGCCTGGTTAGTCTGCGCTATCTGACATGGCGCTTAACCGAGACCCTGAGCTTC

bcsA_ori 301 GATACATGGTTGCAGGGTCTGCTTGGGACAATGCTTCTGGTGGCGGAACTTTACGCCCTG
bcsA_opt 301 GATACCTGGCTGCAGGGCCTGCTGGGTACAATGCTGCTGGTTGCAGAACTGTACGCCCTG

bcsA_ori 361 ATGATGCTGTTCCTCAGCTATTTCCAGACGATCGCGCCATTGCATCGTGCGCCTCTGCCG
bcsA_opt 361 ATGATGCTGTTCCTGAGCTACTTCCAAACCATCGCCCCGTTACATCGTGCCCCGTTACCG

bcsA_ori 421 CTGCCGCCGAACCCTGACGAATGGCCCACGGTCGATATCTTCGTCCCGACCTACAACGAA
bcsA_opt 421 CTGCCGCCGAACCCTGACGAATGGCCCACGGTCGATATCTTCGTCCCGACCTACAACGAA

bcsA_ori 481 GAATGTGGCGCGAATTATATCGCCCGCCCGACGAACGAACATGCAAAGGCCGGTAATCTT
bcsA_opt 481 GAATGCGGCGCAAATTACATCGCACGTCCTACCAACGAACACGCCAAGGCCGGCAACTTA

bcsA_ori 541 AAGGTTGGGGTCTCATATCCTTATGAGCAGTGCTGCTGTCGGTACATCTTACGCGGCCTGTT
bcsA_opt 541 AAGGTTGGGGTCTCATATCCTTATGAGCAGTGCTGCTGTCGGTACATCTTACGCGGCCTGTT

bcsA_ori 601 GAATGCGGCGCAATTATACATCGCGCAGCGACGGCAGCAACTGCAAAGGCCGGTAACTTT
bcsA_opt 601 GAATGCGGCGCAATTATACATCGCGCAGCGACGGCAGCAACTGCAAAGGCCGGTAACTTT

bcsA_ori 661 AACTACCAACAGCTTTTAATCTCCCATCCCGCATCGTCAACATGCGGTGAGAGGCCGCAAAT
bcsA_opt 661 AACTACCAACAGCTTTTAATCTCCCATCCCGCATCGTCAACATGCGGTGAGAGGCCGCAAAT

bcsA_ori 721 CGAGGCGCGCCACATCCCGTACCGTGACCGCGGCTTGCAGAGCCGCAAGTTCCGCGCAG
bcsA_opt 721 CGAGGCGCGCCACATCCCGTACCGTGACCGCGGCTTGCAGAGCCGCAAGTTCCGCGCAG

bcsA_ori 781 GTGATGCAGACCCCGCATCACTTCTATTCCCCGCCGCCGCTGGTGATGCTGATGGGAGG
bcsA_opt 781 GTGATGCAGACCCCGCATCACTTCTATTCCCCGCCGCCGCTGGTGATGCTGATGGGAGG

297
d. Sequence alignments of original and optimised bcsB:

76.0% identity in 2430 residues overlap; Score: 10934.0; Gap frequency: 0.0%
e. Sequence alignments of original and optimised bcsC:

76.9% identity in 3910 residues overlap; Score: 17885.0; Gap frequency: 0.1%
Sequence alignments of original and optimised bcsD:

81.1% identity in 471 residues overlap; Score: 2264.0; Gap frequency: 0.0%