Towards the Development of a Genome Ablation Strategy for Synthetic Biology

A thesis submitted for the degree of

Doctor of Philosophy

in

Synthetic Biology

by

Rosalía Paula Cardós Elena

Department of Biochemical Engineering
University College London
2021
Declaration

I, Rosalía Paula Cardós Elena, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

A major goal of synthetic biology remains to create a fully orthogonal chassis, co-existing with its environment and unable to interfere with it, which would lead to the realisation of real-world applications of this discipline. Current biocontainment strategies often rely on easily circumvented auxotrophy or suicide networks triggered outside of a controlled environment. Rather than establishing a new genetic network to induce lethality, the present work intends to develop a method called genome ablation to obtain a DNA-free chassis unable to self-replicate or transfer genetic material to other species while remaining biochemically functional. In order to achieve absolute DNA degradation in vivo, exonuclease III and T5 bacteriophage exonuclease were overexpressed in Escherichia coli CSR603, a strain deficient in DNA repair, leading to DNA degradation without filamentation. Antibiotics mitomycin C and ciprofloxacin induced lethality and enhanced DNA degradation. Expression of restriction enzymes PvuII or HpaII, DNA gyrase poison ccdB, or T4 bacteriophage endonuclease denA alongside exonucleases reduced DNA content and cell viability; the lowest DNA content was achieved with E. coli CSR603 expressing PvuII at 4 h post-induction. Nuclease induction was found to exert a high selective pressure, leading to loss of nuclease activity. UV irradiation for 15 min induced lethality and absolute DNA degradation. However, this also induced extensive photodamage and compromised membrane integrity. Despite a functional chassis was not attained and further analysis is required, this is the first demonstration of DNA degradation in vivo applied to chassis engineering.
Impact Statement

There is a myriad of potential applications of synthetic biology in fields like vaccine development, bioremediation, biosensors, biomedicine and gene therapy, but the inexistence of a fully orthogonal microorganism hinders translation of these proofs of concept into products. This thesis presents the novel concept of genome ablation, *in vivo* DNA degradation with the aim of obtaining a DNA-free chassis that is still capable of protein production and maintains metabolic activity. This chassis would have an enormous potential to benefit society: it could contain proteins for bioremediation prior to genome ablation, for instance, and then be released freely to the environment without risk of uncontrolled replication or horizontal gene transfer.

This would have implications at a legislative level: current guidelines on release of genetically engineered microorganisms (GMOs) such as European directive 2001/18/EC on the deliberate release into the environment of genetically modified organisms (2001) define an organism as “any biological entity capable of replication or of transferring genetic material”. This chassis would therefore no longer be considered an organism, requiring consideration by legislation and regulations.

Moreover, most works found in the literature employing nucleases in lethality switches to induce cell death do not quantify DNA degradation *in vivo*. This work may instigate those research groups to reassess their genetic circuits in terms of efficiency of DNA degradation for a genome ablation strategy. In addition, it is the first project related to biocontainment in the department of Biochemical Engineering at UCL, and this work highlights useful considerations to take for future projects such as the impact of the type of media used on reproducibility and the limitations of DNA quantification methods.
Acknowledgements

First and foremost, I would like to thank my PhD supervisor Professor John Ward for his advice and guidance throughout this project. I would also like to thank my secondary supervisor, Dr Brenda Parker, for valuable discussions at the beginning of this project. In addition, I would like to thank the EPSRC for funding this project.

I would like to thank everyone in the Ward lab who has supported me and given me advice, especially Sebastian Schulz, who guided my baby steps in molecular biology.

I have been incredibly fortunate to have met so many amazing people in the department with whom I’ve shared my PhD journey these past years. A special thank you to all past and present members of lab 4.02, particularly Sebastian, Fang, Tania and Chili, who made long lab hours and troubleshooting sessions a little bit more bearable. Many thanks to Tom Folliard, who introduced me to recombinase expression vectors.

A special thank you to everyone in vineyard who has helped me persevere and made me realise countless times that despite things might not be going well in the lab everything was going to be OK: Neha, Alex, Becci, Vincent, Lourdes, Víctor, Mau, Luba, Michael, Matt, Charnett, Noelle, Victoria, Artemis, Matt, and Pedro. I cherish our memories together and I’ll miss the vineyard family.

Thank you to my family: my brothers, Javier and José María, for always offering their unconditional support and being the best older brothers I could ever hope for, and of course my parents, José Julián and María del Carmen, for their unwavering support throughout my journey through higher education, believing in me and in turn making me believe I can do anything I set my mind to.

And lastly, my infinite gratitude to Stefan, to whom I owe everything I have achieved in these past years. Thank you for your love, support and endless patience, I wouldn’t have made it this far without you. I can’t wait to see my thesis next to yours in our bookshelf!
Table of contents

Declaration ........................................................................................................................................... 1
Abstract ............................................................................................................................................... 2
Impact Statement ............................................................................................................................... 3
Acknowledgements ............................................................................................................................. 4
List of figures ....................................................................................................................................... 12
List of tables ....................................................................................................................................... 22
List of abbreviations ........................................................................................................................... 23

Chapter 1. Introduction ...................................................................................................................... 25
1.1 The evolution of synthetic biology ............................................................................................... 25
1.2 Genetic engineering tools .............................................................................................................. 28
  1.2.1 Restriction enzyme cloning .................................................................................................... 28
  1.2.2 Ligation-independent cloning .............................................................................................. 28
  1.2.3 PCR-based cloning ............................................................................................................... 29
  1.2.4 Recombinatorial cloning ....................................................................................................... 31
1.3 The potential of synthetic biology ............................................................................................... 32
  1.3.1 Therapeutic applications ...................................................................................................... 33
  1.3.2 Applications in metabolic engineering ............................................................................... 34
  1.3.3 Applications in bioremediation ............................................................................................ 34
1.4 Orthogonality and biocontainment in synthetic biology ............................................................. 35
  1.4.1 Biocontainment via auxotrophy ........................................................................................... 36
  1.4.2 Biocontainment via induced lethality: kill switches .............................................................. 37
  1.4.3 Biocontainment via addiction mechanisms ......................................................................... 37
  1.4.4 Biocontainment via xenobiology ....................................................................................... 38
  1.4.5 DNA-free chassis .................................................................................................................. 39
    1.4.5.1 Minicells ...................................................................................................................... 40
    1.4.5.2 Maxicells .................................................................................................................... 42
  1.4.6 Resistance to biocontainment measures .............................................................................. 43
1.5 DNA damage in bacteria ............................................................................................................. 44
1.5.1 Double-strand breaks (DSBs) ................................................................. 45
1.5.2 DNA damage inducing agents ............................................................... 46
  1.5.2.1 Endonucleases .............................................................................. 46
  1.5.2.2 CRISPR-Cas ................................................................................. 47
  1.5.2.3 Reactive oxygen species .............................................................. 47
  1.5.2.4 UV damage ................................................................................ 48
  1.5.2.5 Bacteriocins ................................................................................ 48
  1.5.2.6 DNA gyrase inhibitors ................................................................. 49
  1.5.2.7 Radiomimetic antibiotics ............................................................. 49
  1.5.2.8 Alkylating agents ....................................................................... 50
1.5.3 DNA repair in prokaryotes ................................................................. 50
  1.5.3.1 Photoreactivation ...................................................................... 50
  1.5.3.2 Base excision repair (BER) ......................................................... 50
  1.5.3.3 Nucleotide excision repair (NER) .............................................. 51
  1.5.3.4 SOS response ............................................................................ 51
  1.5.3.5 Recombination repair ................................................................. 52
1.6 Aims and objectives ........................................................................... 53

Chapter 2. Materials and Methods ............................................................ 54
2.1 Materials ............................................................................................... 54
  2.1.1 Strains ............................................................................................. 54
  2.1.2 Plasmid DNA ................................................................................ 56
  2.1.3 Chemicals ....................................................................................... 56
  2.1.4 Oligonucleotides .......................................................................... 56
  2.1.5 gBlocks .......................................................................................... 57
  2.1.6 Antibiotic stocks ........................................................................... 57
  2.1.7 Growth media ................................................................................ 57
2.2 Methods ............................................................................................... 58
  2.2.1 DNA manipulation and analysis .................................................... 58
    2.2.1.1 Sequencing ................................................................................ 58
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1.2</td>
<td>DNA preparation for cloning</td>
</tr>
<tr>
<td>2.2.1.3</td>
<td>Restriction enzyme digestions</td>
</tr>
<tr>
<td>2.2.1.4</td>
<td>Agarose gel electrophoresis</td>
</tr>
<tr>
<td>2.2.1.5</td>
<td>PCR</td>
</tr>
<tr>
<td>2.2.1.6</td>
<td>Colony PCR</td>
</tr>
<tr>
<td>2.2.1.7</td>
<td>Ligation</td>
</tr>
<tr>
<td>2.2.1.8</td>
<td>Circular Polymerase Extension Cloning (CPEC)</td>
</tr>
<tr>
<td>2.2.1.9</td>
<td>Gibson assembly</td>
</tr>
<tr>
<td>2.2.1.10</td>
<td>Total DNA extraction</td>
</tr>
<tr>
<td>2.2.1.11</td>
<td>Field-Inversion Gel Electrophoresis (FIGE)</td>
</tr>
<tr>
<td>2.2.1.12</td>
<td>PicoGreen assay</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Preparation of chemically competent cells</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Preparation of electrocompetent cells</td>
</tr>
<tr>
<td>2.2.4</td>
<td>Chemical transformation</td>
</tr>
<tr>
<td>2.2.5</td>
<td>Electroporation</td>
</tr>
<tr>
<td>2.2.6</td>
<td>Plasmid design and assembly</td>
</tr>
<tr>
<td>2.2.6.1</td>
<td>pQR1866, pQR1867, pQR1868</td>
</tr>
<tr>
<td>2.2.6.2</td>
<td>pQR1869, pQR1870</td>
</tr>
<tr>
<td>2.2.6.3</td>
<td>Cloning restriction enzymes in pTTQ18</td>
</tr>
<tr>
<td>2.2.6.4</td>
<td>Cloning restriction enzymes in pQR445</td>
</tr>
<tr>
<td>2.2.6.5</td>
<td>Cloning methyltransferases in pACYC184</td>
</tr>
<tr>
<td>2.2.6.6</td>
<td>pPVU1t and pPVU1t2</td>
</tr>
<tr>
<td>2.2.6.7</td>
<td>pQR1871</td>
</tr>
<tr>
<td>2.2.6.8</td>
<td>pQR1872, pQR1873, pQR1874, pQR1875</td>
</tr>
<tr>
<td>2.2.6.9</td>
<td>pQR1876</td>
</tr>
<tr>
<td>2.2.6.10</td>
<td>pQR1877, pQR1878, pQR1879, pQR1880</td>
</tr>
<tr>
<td>2.2.6.11</td>
<td>pQR1886</td>
</tr>
<tr>
<td>2.2.7</td>
<td>Colony screening</td>
</tr>
<tr>
<td>2.2.7.1</td>
<td>Catechol 2,3-dioxygenase screening</td>
</tr>
</tbody>
</table>
2.2.7.2 CV2025 transaminase (TAm) screening ........................................... 71

2.3 Protein expression studies ........................................................................ 71
  2.3.1 Protein extract preparation ..................................................................... 71
  2.3.2 Protein activity assays ........................................................................... 71
    2.3.2.1 Exonuclease activity assays ............................................................. 71
    2.3.2.2 Endonuclease activity assays ........................................................... 71
    2.3.2.3 Transaminase assay ........................................................................... 72
    2.3.2.4 Meta pathway and catechol 2,3-dioxygenase assays ......................... 72
  2.3.3 SDS-PAGE .................................................................................................. 73
  2.3.4 Western blot for His-tagged protein detection ........................................ 73

2.4 Experiments involving cell growth and analysis .................................... 74
  2.4.1 Working cell banks for exonuclease expression growth curves ............... 74
  2.4.2 Growth curves for exonuclease overexpression ...................................... 74
  2.4.3 Determination of growth-inhibiting MMC and CI concentrations .......... 74
  2.4.4 Combined antibiotic addition and nuclease induction experiments .......... 75
  2.4.5 DSB-inducing protein induction experiments ......................................... 75
  2.4.6 Serial dilutions and colony-forming unit (cfu) determination .................. 76
  2.4.7 Monitoring cell growth and fluorescence in a 96 well plate .................... 76
  2.4.8 Genome ablation and UV irradiation experiments ................................ 77
  2.4.9 Cell staining ............................................................................................. 78
  2.4.10 Flow cytometry ....................................................................................... 78
  2.4.11 Fluorescence microscopy ....................................................................... 78

Chapter 3. Effect of exonuclease overexpression in E. coli CSR603 ............. 80
  3.1 Introduction .................................................................................................. 80
  3.2 Results and discussion .............................................................................. 82
    3.2.1 Cloning single exonucleases in pTTQ18 .............................................. 82
    3.2.2 Cloning exonuclease pairs in pTTQ18 ................................................. 87
    3.2.3 Exonuclease expression ......................................................................... 88
    3.2.3.1 SDS PAGE .......................................................................................... 88
3.2.3.2 Exonuclease activity in vitro ........................................................................90
3.2.4 Effect of exonuclease overexpression on growth ........................................91
3.2.5 Visualisation of exonuclease-expressing E. coli CSR603 with a widefield
    microscope ...........................................................................................................96
3.2.6 Effect of exonuclease expression on total DNA content ..............................99
3.2.7 Loss of viability in cells expressing exonucleases ........................................104
3.2.8 *In vitro* exonuclease activity assays ............................................................105

3.3 Conclusions ......................................................................................................116

Chapter 4. Use of DSB-inducing antibiotics with nuclease expression .............117
4.1 Introduction .......................................................................................................117
4.2 Results .............................................................................................................118
    4.2.1 Evaluating different concentrations of MMC and CI ............................118
    4.2.2 Combining drug addition with exonuclease overexpression in E. coli
        CSR603 pQR1870 ......................................................................................120
    4.2.3 Combining MMC and CI with endo- and exonuclease expression ......124
        4.2.3.1 CSR603 .........................................................................................124
        4.2.3.2 DH5αZ1 .........................................................................................126
    4.2.4 Comparison of all strategies ..................................................................131
4.3 Discussion and conclusions ...........................................................................134

Chapter 5. DNA-damaging protein expression as a DSB-inducing agent ........136
5.1 Introduction .......................................................................................................136
5.2 Results and discussion ...................................................................................138
    5.2.1 Unsuccessful attempts at cloning DSB-inducing proteins .................138
        5.2.1.1 Cloning under the Ptac promoter in pTTQ18 and pQR445 .........138
        5.2.1.2 Using methyltransferases: pACYC184 and pPVU1t .................141
    5.2.2 Using a dual recombinase expression system ........................................147
        5.2.2.1 Testing the sensitivity of the dual recombinase expression system.148
        5.2.2.2 Cloning DSB-inducing proteins using the dual recombinase expression
            system ....................................................................................................149
5.2.2.3 Cloning the exonuclease cassette into the dual recombinase expression system .................................................................160
5.2.3 Ascertaining the optimal induction time ........................................161
5.2.4 Testing different DSB-inducing proteins .......................................170
5.2.5 Comparing DH5αZ1 and CSR603 .............................................175
5.2.6 Addressing the increase in viability post-induction .....................179
  5.2.6.1 Testing restriction enzyme activity .......................................179
  5.2.6.2 Verifying the presence of the controller and expression plasmids ....182
  5.2.6.3 Investigating plasmid presence post-induction .....................186
5.3 Summary and conclusions ..................................................................189

Chapter 6. Using UV radiation to achieve genome ablation and assessment of a genome-ablated chassis .........................................................193
6.1 Introduction .......................................................................................193
6.2 Results and discussion ......................................................................194
  6.2.1 Use of M9 to enhance reproducibility ........................................194
  6.2.2 UV irradiation as an additional DSB-inducing agent ..............197
  6.2.3 Cloning constitutive expression cassettes into pQR1871 ..........200
  6.2.4 Assessing protein activity post-genome ablation ...................203
    6.2.4.1 PvuII and exonuclease activities ......................................204
    6.2.4.2 sfGFP .......................................................................205
    6.2.4.3 Pseudomonas meta pathway ..........................................207
    6.2.4.4 C. violaceum CV2025 TAm ...........................................210
  6.2.5 GA cells as recipients for transformation ..................................214
  6.3 Flow cytometric analysis of GA cells ..........................................218
  6.3 Conclusions ...................................................................................226

Chapter 7. Conclusions and future work ......................................................229
7.1 Summary and conclusions ...............................................................229
7.2 Future work .....................................................................................231

Appendix .................................................................................................234
8.1 Plasmid maps ....................................................................................234
8.2 Oligonucleotide list ........................................................................................................... 262
8.3 PCR conditions .................................................................................................................. 267
8.4 gBlock sequences .............................................................................................................. 274
8.5 Exonuclease activity assay gels .......................................................................................... 277
8.6 Total DNA extraction gels .................................................................................................. 303

Bibliography ............................................................................................................................ 307
List of figures

Figure 1-1. Timeline illustrating the main achievements in synthetic biology over the past decades. Adapted from Cameron et al. (2014). ................................................................. 27
Figure 1-2. Depiction of how CPEC works, starting from a restriction digest- or PCR-prepared vector and an insert with overlaps inserted via PCR. ........................................ 30
Figure 1-3. Illustration depicting how Gibson assembly works ........................................ 31
Figure 1-4. Illustration of minicell formation through abnormal cell division.................. 41
Figure 3-1. Simplified plasmid map of the different exonuclease plasmids .................... 81
Figure 3-2. Gel depicting PCR amplification of exonuclease III from the E. coli MG1655 genome .................................................................................................................. 82
Figure 3-3. Gel depicting PCR amplification of λ exonuclease from the λ bacteriophage genome .................................................................................................................. 82
Figure 3-4. Gel depicting PCR amplification of T5 exonuclease from the T5 bacteriophage genome ............................................................................................................. 83
Figure 3-5. Illustration of the two-cistron system created in order to clone the exonucleases into vector pTTQ18 .......................................................................................... 84
Figure 3-6. Gel depicting the CPEC reaction for assembly of the exonuclease III insert into the pTTQ18 backbone ....................................................................................... 85
Figure 3-7. Gel depicting colony PCR after CPEC assembling pQR1866 ....................... 86
Figure 3-8. Gel depicting colony PCR after CPEC assembling pQR1868 and pQR1867 .............................................................................................................................. 86
Figure 3-9. Gel depicting PCR amplification of the inserts needed for pQR1869 and pQR1870 ...................................................................................................................... 87
Figure 3-10. Gel depicting colony PCR after CPEC assembling pQR1869 and pQR1870 ....................................................................................................................... 88
Figure 3-11. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1866 ............... 89
Figure 3-12. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1867 ............... 89
Figure 3-13. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1868 ............... 89
Figure 3-14. Activity assay gels for E. coli CSR603 with pQR1866 (exonuclease III), pQR1867 (λ exonuclease) and pQR1868 (T5 exonuclease) ......................................... 90
Figure 3-15. Growth curves of E. coli CSR603 harbouring the different exonuclease plasmids .................................................................................................................... 92
Figure 3-16. Side-to-side comparison of the E. coli CSR603 pQR1868 growth curves 95
Figure 3-17. Brightfield microscope images of E. coli CSR603 pQR1866 .................... 97
Figure 3-18. Brightfield microscope images of E. coli CSR603 pQR1868 .................... 98
Figure 3-19. Total DNA quantification of different E. coli CSR603 cultures expressing exonucleases .......................................................................................................................... 100
Figure 3-20. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on EcoRV-digested λ DNA ................................................................................................................. 106
Figure 3-21. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on Bpml-digested λ DNA .................................................................................................................. 106
Figure 3-22. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on HindIII-digested λ DNA .................................................................................................................. 107
Figure 3-23. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on EcoRV-digested λ DNA ................................................................................................................. 107
Figure 3-24. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on Bpml-digested λ DNA .................................................................................................................. 108
Figure 3-25. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on HindIII-digested λ DNA .................................................................................................................. 108
Figure 3-26. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on EcoRV-digested λ DNA ................................................................................................................. 109
Figure 3-27. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on Bpml-digested λ DNA .................................................................................................................. 109
Figure 3-28. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on HindIII-digested λ DNA .................................................................................................................. 110
Figure 3-29. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on EcoRV-digested λ DNA ................................................................................................................. 110
Figure 3-30. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on Bpml-digested λ DNA .................................................................................................................. 111
Figure 3-31. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on HindIII-digested λ DNA .................................................................................................................. 111
Figure 3-32. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 on EcoRV-digested λ DNA ................................................................................................................. 112
Figure 3-33. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 on Bpml-digested λ DNA .................................................................................................................. 112
Figure 3-34. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 on HindIII-digested λ DNA .................................................................................................................. 113
Figure 4-1. Change in OD600 post-drug addition in E. coli CSR603 pQR1870 ........... 119
Figure 4-2. Growth curves of E. coli CSR603 pQR1870 at different MMC and CI concentrations ............................................................................................................................... 121
Figure 4-3. Densitometric analysis of total DNA content in *E. coli* CSR603 with or without exonuclease expression and at different MMC, CI concentrations ........................................... 122
Figure 4-4. DNA gels with samples from different *E. coli* CSR603 pQR1871 pQR1872 cultures to which exonuclease expression was induced and MMC and Cl added ...... 123
Figure 4-5. Growth curves of *E. coli* CSR603 pQR1871 pQR1872, either non-induced (n/I) or induced (Ind.), at different MMC and Cl concentrations.......................................................... 125
Figure 4-6. Densitometric analysis of total DNA content in *E. coli* CSR603 pQR1871 pQR1872 with or without Pvull and exonuclease expression and at different MMC, Cl concentrations .............................................................................. 126
Figure 4-7. Change in OD$_{600}$ post-drug addition in *E. coli* DH5αZ1 pQR1871 pQR1872 .......................................................................................................................... 128
Figure 4-8. Growth curves of *E. coli* DH5αZ1 pQR1871 pQR1872, either non-induced (n/I) or induced (Ind.), at different MMC and Cl concentrations..................................................... 130
Figure 4-9. Densitometric analysis of total DNA content in *E. coli* DH5αZ1 pQR1871 pQR1872 with or without Pvull and exonuclease expression and at different MMC, Cl concentrations .............................................................................. 131
Figure 4-10. Comparison between the different induced samples for the three main experiments in this chapter ....................................................................................................... 132
Figure 5-1. PCR amplification of the *NotI* and *HpaII* genes for cloning via CPEC into pTTQ18 ...................................................................................................................... 138
Figure 5-2. PCR amplification of the *BamHI* gene for cloning via CPEC into pTTQ18 ...................................................................................................................... 139
Figure 5-3. PCR amplification of the *Pvull* gene for cloning via CPEC into pTTQ18 .. 139
Figure 5-4. PCR amplification of the *Pvull* CDS for cloning via CPEC in pQR445 ..... 140
Figure 5-5. PCR amplification of the NotI and the Pvull methyltransferase genes ..... 142
Figure 5-6. PCR amplification of the BamHI methyltransferase gene ....................... 142
Figure 5-7. PCR amplification of the Hpal methyltransferase gene ....................... 143
Figure 5-8. FIGE gel depicting the methyltransferase activity test in vivo in HB101 ... 144
Figure 5-9. Gel with 50 ng of the EcoRI-digested pPVU1 plasmid ......................... 145
Figure 5-10. PCR amplification of pPVU1t for linearization ..................................... 145
Figure 5-11. pPVU1t2 verification via plasmid digestion .......................................... 146
Figure 5-12. FIGE to test methyltransferase activity test in vivo for pPVU1t ............ 146
Figure 5-13. Depiction of integrase-mediated expression in the dual plasmid system. .......................................................................................................................... 148
Figure 5-14. Graph depicting relative fluorescence units (RFU) normalised by OD\textsubscript{600} in E. coli DH5αZ1 harbouring the dual recombinase system ................................................. 148
Figure 5-15. Scheme depicting the cloning strategy for DSB-inducing proteins into the controller plasmid of the dual recombinase system ....................................................... 150
Figure 5-16. PCR of the expression vector backbone introducing BamHI and XhoI sites. ........................................................................................................................................ 151
Figure 5-17. PCR of the PvuII CDS ...................................................................................... 151
Figure 5-18. Screening of expression plasmids containing the PvuII CDS via miniprep digestion with BamHI and XhoI .......................................................................................... 152
Figure 5-19. In vitro activity test of PvuII in E. coli DH5αZ1 harbouring the controller plasmid and pQR1872 ........................................................................................................ 152
Figure 5-20. Colony PCR screening verifying CDS insertion in the expression plasmid ................................................................................................................................. 153
Figure 5-21. Gel depicting 50 ng of the different inserts ready to clone into the expression plasmid .......................................................................................................................... 154
Figure 5-22. Miniprep digestions of the colonies belonging to lanes 10, 11, 15, 21 and 25 screened in Figure 5-20 ........................................................................................................ 154
Figure 5-23. In vitro activity test of HpaII in E. coli DH5αZ1 harbouring the controller plasmid and pQR1873 ........................................................................................................ 155
Figure 5-24. Snapshot of SnapGene showing a portion of the alignment of the original designed sequence of the beginning of the PvuII CDS in pQR1872 .................................. 157
Figure 5-25. Snapshot of SnapGene depicting the whole PvuII gene sequence and highlighting the other ORF found within the sequenced PvuII CDS in pQR1872 ........ 158
Figure 5-26. Snapshot of SnapGene showing a portion of the alignment of the original designed sequence of the beginning of the ccdB CDS in pQR1875 ...................... 159
Figure 5-27. PCR amplification of the exonuclease cassette prior to cloning into the controller plasmid ................................................................................................................... 160
Figure 5-28. Colony screening of pQR1871 ........................................................................ 161
Figure 5-29. Growth curves of DH5αZ1 pQR1871 pQR1872 at the different induction conditions .............................................................................................................................. 163
Figure 5-30. The graph shows the escape frequencies measured at different times after inoculation for the different induction conditions .......................................................... 165
Figure 5-31. Variation in DNA content of samples with different induction conditions in which the first induction occurs at 3 h or at 7 h ........................................................................ 168
Figure 5-32. Results from Figure 5-31 (left) without the non-induced control ............ 169
Figure 5-33. Growth curves of E. coli DH5αZ1 expressing exonucleases (pQR1871) together with a DSB-inducing protein ................................................................. 171
Figure 5-34. Ratio of the viable cell count obtained at 7 h and 24 h after inoculation divided by the viable count at the time of induction ........................................... 172
Figure 5-35. Variation in DNA content of E. coli DH5αZ1 expressing exonucleases and DSB-inducing proteins ................................................................. 173
Figure 5-36. Variation in DNA content displayed in Figure 5-35 without the control of E. coli DH5αZ1 expressing exonucleases and DSB-inducing proteins ......................... 174
Figure 5-37. Growth curves of induced DH5αZ1 pQR1871 pQR1873 and CSR603 pQR1871 pQR1872 ................................................................. 176
Figure 5-38. Ratio of the viable cell count obtained for DH5αZ1 pQR1871 pQR1873 and CSR603 pQR1871 pQR1872 ................................................................. 177
Figure 5-39. Variation in DNA content of induced E. coli DH5αZ1 pQR1871 pQR1873 and CSR603 pQR1871 pQR1872 ................................................................. 178
Figure 5-40. Variation in DNA content displayed in Figure 5-39 for induced E. coli DH5αZ1 pQR1871 pQR1873 and CSR603 pQR1871 pQR1872 without their respective controls ................................................................. 178
Figure 5-41. In vitro restriction enzyme activity tests for E. coli DH5αZ1 pQR1871 pQR1872, E. coli DH5αZ1 pQR1871 pQR1873 and E. coli CSR603 pQR1871 pQR1872 ................................................................. 180
Figure 5-42. In vitro restriction enzyme activity tests for E. coli DH5αZ1 pQR1871 pQR1873, E. coli DH5αZ1 pQR1871 pQR1872 and E. coli CSR603 pQR1871 pQR1872 ................................................................. 181
Figure 5-43. Miniprep digestions with XhoI and BamHI of DH5αZ1 pQR1871 pQR1872 cultures induced and grown overnight ......................................................... 183
Figure 5-44. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1873 cultures induced and grown overnight ......................................................... 183
Figure 5-45. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1874 cultures induced and grown overnight ......................................................... 184
Figure 5-46. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1875 cultures induced and grown overnight ......................................................... 184
Figure 5-47. Miniprep digestions XhoI and BamHI of CSR603 pQR1871 pQR1872 cultures induced and grown overnight ......................................................... 185
Figure 5-47. Miniprep digestions XhoI and BamHI of CSR603 pQR1871 pQR1872 colonies grown in M9 Amp Cam at different induction conditions ......................................................... 196
Figure 6-2. FIGE image of UV-irradiated E. coli CSR603 pQR1871 pQR1872 irradiated at different times .................................................................................................................. 198
Figure 6-3. FIGE gel of E. coli CSR603 pQR1871 pQR1872 induced samples, irradiated at 15 min ........................................................................................................................................ 199
Figure 6-4. pACYC184 digestion with Ncol and EcoRI .............................................................. 201
Figure 6-5. PCR amplification of the pQR1877 backbone inserting Smal and a BsrGI sites ........................................................................................................................................... 201
Figure 6-6. PCR amplification of the catechol pathway sequence prior to digestion and ligation into the pQR1877 backbone .................................................................................................................. 202
Figure 6-7. PCR amplification of the CV2025 TAm and tinsel purple prior to digestion and ligation into the pQR1877 backbone ...................................................................................................... 202
Figure 6-8. Gel depicting in vitro activity assays for E. coli CSR603 pQR1871 pQR1872 subjected to the UV irradiation protocol .................................................................................................................. 204
Figure 6-9. Pictures of the pellets of E. coli CSR603 pQR1877 pQR1872, non-induced and induced ........................................................................................................................................ 206
Figure 6-10. Graph depicting the relative fluorescence units (RFU) normalised by OD600 of non-induced and induced E. coli CSR603 pQR1872 pQR1877 subjected to different UV irradiation times ........................................................................................................................................ 206
Figure 6-11. Scheme depicting conversion of benzoate through to 2-hydroxymuconic aldehyde via the Pseudomonas upper pathway........................................................................................................... 208
Figure 6-12. Graph depicting the results of the catechol 2,3-dioxygenase assay on non-induced and induced E. coli CSR603 pQR1872 pQR1878 subjected to different UV irradiation times ........................................................................................................... 209
Figure 6-13. Western blot of UV-irradiated E. coli CSR603 pQR1886 pQR1872, non-induced and induced ........................................................................................................................................ 213
Figure 6-14. Flow cytometry analysis of GA transformants with an RFP plasmid .......... 216
Figure 6-15 Flow cytometry analysis of genome-ablated (GA) E. coli CSR603 pQR1871 pQR1872 ........................................................................................................................................... 220
Figure 6-16. Widefield microscope images of non-induced E. coli CSR603 pQR1871 pQR1872 cells ........................................................................................................................................ 222
Figure 6-17. Widefield microscope images of induced E. coli CSR603 pQR1871 pQR1872 cells ........................................................................................................................................ 223
Figure 8-1. pTTQ18 plasmid map extracted from SnapGene .................................................. 234
Figure 8-2. pQR445 plasmid map extracted from SnapGene ................................................ 235
Figure 8-3. pACYC184 plasmid map extracted from SnapGene .............................................. 236
Figure 8-4. pQR1866 map extracted from SnapGene ................................................................. 237
Figure 8-34. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 2) on HindIII-digested λ DNA ................................................................. 279
Figure 8-35. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on EcoRV-digested λ DNA .................................................. 280
Figure 8-36. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on BpmI-digested λ DNA ............................................. 280
Figure 8-37. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on HindIII-digested λ DNA ............................................. 281
Figure 8-38. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on EcoRV-digested λ DNA ............................................. 281
Figure 8-39. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on BpmI-digested λ DNA ............................................. 282
Figure 8-40. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on HindIII-digested λ DNA ............................................. 282
Figure 8-41. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on EcoRV-digested λ DNA ............................................. 283
Figure 8-42. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on BpmI-digested λ DNA ............................................. 283
Figure 8-43. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on HindIII-digested λ DNA ............................................. 284
Figure 8-44. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on EcoRV-digested λ DNA ............................................. 284
Figure 8-45. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on BpmI-digested λ DNA ............................................. 285
Figure 8-46. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on HindIII-digested λ DNA ............................................. 285
Figure 8-47. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on EcoRV-digested λ DNA ............................................. 286
Figure 8-48. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on BpmI-digested λ DNA ............................................. 286
Figure 8-49. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on HindIII-digested λ .................................................. 287
Figure 8-50. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 (replicate 2) on EcoRV-digested λ DNA ............................................. 287
Figure 8-51. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 (replicate 2) on BpmI-digested λ DNA ............................................. 288
Figure 8-52. Activity tests of protein extracts from induced E. coli CSR603 pQR1867
(replicate 2) on HindIII-digested λ DNA ................................................................. 288
Figure 8-53. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 1) on EcoRV-digested λ DNA ................................................................. 289
Figure 8-54. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 1) on BpmI-digested λ DNA ................................................................. 289
Figure 8-55. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 1) on HindIII-digested λ DNA ................................................................. 290
Figure 8-56. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 2) on EcoRV-digested λ DNA ................................................................. 290
Figure 8-57. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 2) on BpmI-digested λ DNA ................................................................. 291
Figure 8-58. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868
(replicate 2) on HindIII-digested λ DNA ................................................................. 291
Figure 8-59. Activity tests of protein extracts from induced E. coli CSR603 pQR1868
(replicate 2) on EcoRV-digested λ DNA ................................................................. 292
Figure 8-60. Activity tests of protein extracts from induced E. coli CSR603 pQR1868
(replicate 2) on BpmI-digested λ DNA ................................................................. 292
Figure 8-61. Activity tests of protein extracts from induced E. coli CSR603 pQR1868
(replicate 2) on HindIII-digested λ DNA ................................................................. 293
Figure 8-62. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 1) on EcoRV-digested λ DNA ................................................................. 293
Figure 8-63. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 1) on BpmI-digested λ DNA ................................................................. 294
Figure 8-64. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 1) on HindIII-digested λ DNA ................................................................. 294
Figure 8-65. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 2) on EcoRV-digested λ DNA ................................................................. 295
Figure 8-66. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 2) on BpmI-digested λ DNA ................................................................. 295
Figure 8-67. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869
(replicate 2) on HindIII-digested λ DNA ................................................................. 296
Figure 8-68. Activity tests of protein extracts from induced E. coli CSR603 pQR1869
(replicate 2) on EcoRV-digested λ DNA ................................................................. 296
Figure 8-69. Activity tests of protein extracts from induced E. coli CSR603 pQR1869
(replicate 2) on BpmI-digested λ DNA ................................................................. 297
Figure 8-70. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 (replicate 2) on HindIII-digested λ DNA ................................................................. 297
Figure 8-71. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on EcoRV-digested λ DNA ................................................................. 298
Figure 8-72. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on BpmI-digested λ DNA ................................................................. 298
Figure 8-73. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on HindIII-digested λ DNA ................................................................. 299
Figure 8-74. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on EcoRV-digested λ DNA ................................................................. 299
Figure 8-75. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on BpmI-digested λ DNA ................................................................. 300
Figure 8-76. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on HindIII-digested λ DNA ................................................................. 300
Figure 8-77. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on EcoRV-digested λ DNA ................................................................. 301
Figure 8-78. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on BpmI-digested λ DNA ................................................................. 301
Figure 8-79. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on HindIII-digested λ DNA ................................................................. 302
Figure 8-80. DNA gels with samples from different E. coli CSR603 cultures to which PvuII and exonuclease expression was induced and MMC and CI added .................. 303
Figure 8-81. DNA gels with samples from different E. coli DH5αZ1 cultures to which PvuII and exonuclease expression was induced and MMC and CI added .................. 304
List of tables

Table 1-1. Summary of the main sources of DSBs in bacteria........................................45
Table 2-1. Strains used in this work, together with their genotype ........................................54
Table 2-2. Plasmids used as template DNA, cloning or expression vectors in this work ..............................................................56
Table 2-3. Plasmids constructed in this work........................................................................64
Table 3-1. Maximum OD\textsubscript{600} reached by each non-induced culture of E. coli CSR603 with
the different exonuclease plasmids. ..................................................................................94
Table 3-2. Viable cell count of E. coli CSR603 pQR1868 in permissive (+Amp) or non-
permissive (+Amp, IPTG) conditions ..............................................................................105
Table 4-1. Compilation of all combinations of MMC and CI assayed in Figure 4-1........118
Table 5-1. OD\textsubscript{600} measurements of different DH5aZ1 cultures harbouring the controller
plasmid and the different endonuclease plasmids 4 h after induction ...............................156
Table 5-2. Percentage of different E. coli strains expressing nucleases cells maintaining
either the controller plasmid (Amp), the expression plasmid (Amp), or both (Amp + Cam)
........................................................................................................................................187
Table 6-1. Transaminase assay for UV-irradiated samples performed in triplicates
alongside the controls ........................................................................................................211
Table 6-2. Composition of the different conditions in the TAm assay depicted in Table
6-1.......................................................................................................................................212
Table 6-3. Flow cytometry results compilation of the RFP plasmid transformation
experiment .........................................................................................................................217
Table 6-4. Flow cytometry analysis of genome-ablated (GA) E. coli CSR603 pQR1871
pQR1872 ..........................................................................................................................221
Table 8-1. Oligonucleotides used for sequencing using Eurofins Genomics tube
sequencing service (Ebersberg, Germany). .......................................................................262
Table 8-2. Oligonucleotides used for gene amplification in cloning in Chapter 3........263
Table 8-3. Oligonucleotides used for cloning in chapter 5.................................................264
Table 8-4. Oligonucleotides used for cloning in chapter 6................................................266
Table 8-5. Conditions for every PCR performed in Chapter 3 .......................................267
Table 8-6. Conditions for every PCR performed in Chapter 5 .......................................269
Table 8-7. Conditions for every PCR performed in Chapter 6 .......................................273
Table 8-8. DNA gels with samples from different E. coli DH5aZ1 pQR1871 pQR1872
cultures at different induction conditions ........................................................................305
Table 8-9. DNA gels with samples from E. coli DH5aZ1 and CSR603 in Chapter 5,
sections 5.2.3 to 5.2.5 ......................................................................................................306
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>Bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>Cam</td>
<td>Chloramphenicol</td>
</tr>
<tr>
<td>CDS</td>
<td>Coding sequence</td>
</tr>
<tr>
<td>Ci</td>
<td>Ciprofloxacin</td>
</tr>
<tr>
<td>CPEC</td>
<td>Circular polymerase extension cloning</td>
</tr>
<tr>
<td>dNMP</td>
<td>Deoxyribonucleoside monophosphate</td>
</tr>
<tr>
<td>DSB</td>
<td>Double strand break</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FIGE</td>
<td>Field-inversion gel electrophoresis</td>
</tr>
<tr>
<td>FMO</td>
<td>Fluorescence-minus-one</td>
</tr>
<tr>
<td>GA cells</td>
<td>Genome-ablated cells</td>
</tr>
<tr>
<td>GMO</td>
<td>Genetically modified organism</td>
</tr>
<tr>
<td>His-tag</td>
<td>Hexahistidine tag</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl ß-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>Km</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PFGE</td>
<td>Pulsed-field gel electrophoresis</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>LDS</td>
<td>Lithium dodecyl sulphate</td>
</tr>
<tr>
<td>LIC</td>
<td>Ligation-independent cloning</td>
</tr>
<tr>
<td>MCS</td>
<td>Multiple cloning site</td>
</tr>
<tr>
<td>MFI</td>
<td>Median fluorescence intensity</td>
</tr>
<tr>
<td>MMC</td>
<td>Mitomycin C</td>
</tr>
<tr>
<td>MOPS</td>
<td>3-(N-morpholino)propanesulfonic acid</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>Nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>PLP</td>
<td>Pyridoxal 6-phosphate</td>
</tr>
<tr>
<td>RBS</td>
<td>Ribosome binding site</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate - polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>sfGFP</td>
<td>Superfolder GFP</td>
</tr>
<tr>
<td>SLIC</td>
<td>Sequence and ligation-independent cloning</td>
</tr>
<tr>
<td>SOC</td>
<td>Super optimal broth with catabolite repression</td>
</tr>
<tr>
<td>SSB</td>
<td>Single strand break</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single-stranded DNA</td>
</tr>
<tr>
<td>TAm</td>
<td>Transaminase/aminotransferase</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris borate EDTA</td>
</tr>
<tr>
<td>TBST</td>
<td>Tris-buffered saline with Tween 20</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-EDTA</td>
</tr>
<tr>
<td>Tet</td>
<td>Tetracycline</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>V</td>
<td>Volts</td>
</tr>
<tr>
<td>WCB</td>
<td>Working cell bank</td>
</tr>
</tbody>
</table>
1 Chapter 1. Introduction

1.1 The evolution of synthetic biology

Despite the concept of synthetic biology seeming to have arisen within the past few decades, the idea of this discipline was suggested much longer ago. The first mention of the term ‘synthetic biology’ dates from the early 20th century, in Stéphane Leduc’s _Physicochemical theory of life and spontaneous generation_ (1910). Leduc predicted Biology would experience the same evolution as all other natural sciences, becoming successively descriptive, analytical and synthetic. Later on, genetist Waclaw Szybalski foresaw that synthetic biology research would enable adding control modules to existing modules or even design whole new genomes in a field of unlimited potential (Szybalski, 1974).

It was not until late 20th century that technological innovation allowed for the development of this discipline. Notable milestones are depicted in Figure 1-1. In the 60s it was first hypothesised how gene regulation mechanisms would model cell activity, but research on regulation pathways was limited until the invention of PCR and development of molecular cloning. In the 90s, the availability of computational tools, genome sequencing and high-throughput genomic techniques enabled the generation of the field of systems biology to reverse-engineer cellular networks, treating them as hierarchical functional modules reminiscent of engineering (Cameron et al., 2014).

In contrast to this top-down approach, a bottom-up approach would entail engineering complex networks from molecular building blocks. Translated to the biological field, this implies treating nature as complex interactions, or circuits, of individual modular blocks. Natural gene circuits already exist in nature, such as the circuit that bacteriophage λ utilises to switch between lysis or lysogeny, but the first synthetic cellular networks emerged in the beginning of the 21st century. The repressilator (Elowitz and Leibler, 2000) and genetic toggle switch (Gardner et al., 2000) were the first examples of synthetic networks, employing repressible promoters and fluorescent reporters to
evaluate the behaviour of these circuits. Since then, DNA synthesis has gradually become cheaper and efficient DNA assembly methods such as Gibson assembly enabled rapid and efficient design of synthetic networks, gaining complexity over the years: examples are the development of kill switches (Chan et al., 2015), light-sensing *Escherichia coli* (Levskaya et al., 2005) and use of *E. coli* as biosensor for landmines (Belkin et al., 2017). Other recent milestones include the synthesis of the first eukaryotic chromosome (Annaluru et al., 2014) and design and synthesis of a minimal bacterial genome, with a genome smaller than any other found in nature (Hutchison et al., 2016). More recently, an *E. coli* strain has been completely recoded to utilise 61 codons instead of 64 (Fredens et al., 2019).

Through these advancements, synthetic biology has enabled rapid optimisation of organisms in industrial bioproduction, but there is a myriad of potential applications outside a contained laboratory setting. Despite laboratory strains are unlikely to compete in the environment, it is necessary to minimise the potential risk of genetically modified organism (GMO) interference with the environment by placing genetic safeguards that prevent uncontrolled release, simultaneously addressing public concerns on their safety.
Figure 1-1. Timeline illustrating the main achievements in synthetic biology over the past decades. Adapted from Cameron et al. (2014).
1.2 Genetic engineering tools

1.2.1 Restriction enzyme cloning

Recombinant DNA techniques have evolved significantly since the first DNA manipulation methods in the 1920s. Traditional restriction digestion and ligation for cloning, although reliable and still widely used, is limited by sequence and consequently size of the assembled desired product. In addition, preparation of the different fragments may require methods like gel extraction and multiple purification steps prior to transformation.

Golden Gate assembly is a step ahead from traditional restriction-ligation. It also employs restriction enzymes, but of the IIS type (like *Bsa*I or *Bbs*I), which cleave outside of the recognition site and generate 4-bp compatible overhangs. In Golden Gate, digestion and ligation can be done in one step within 30 min in a reaction with both restriction enzyme and T4 DNA ligase (Engler et al., 2008). However, there is still a dependence on the sequence not containing these restriction sites.

1.2.2 Ligation-independent cloning

Ligation Independent Cloning (LIC) is a cloning method created in the 1990s. Vector and insert are first prepared via PCR or restriction enzyme digestion and designed to have overlapping regions, then mixed in a reaction with T4 DNA polymerase and only one type of dNTP. As well as its polymerase activity, this enzyme also has 3' → 5' exonuclease activity when there are no dNTPs: in presence of only one type of dNTP, T4 DNA polymerase will digest the DNA fragment until it encounters that base, at which point it will add back the dNTP and stall. Then, the complementary overhangs from the vector and different fragments will anneal, resulting in an assembled product with four nicks. The overhangs are designed to be 12 – 15 bp long to provide enough complementarity and stability to ensure they remain annealed upon transformation, when the host DNA repair mechanisms will seal the nicks (Aslanidis and de Jong, 1990).
Nevertheless, LIC also has its drawbacks. Often, only specifically designed plasmids are amenable to LIC given its sequence limitations; therefore, improvements to LIC have been made over the years. Sequence and Ligation Independent Cloning (SLIC) circumvents the need for careful overhang design and relies on homologous recombination instead. Inserts are designed to have 20 bp homology if only one fragment is being cloned, or 40 bp if multiple fragments are required in the cloning strategy. Inserts are generated by either PCR, taking advantage of the fact that in later PCR cycles there might be 5’ overhangs, or by using the exonuclease activity of T4 DNA polymerase. Vector and inserts are combined and annealed, and then recombination is stimulated by the ssDNA overhangs: this can be done more efficiently by adding RecA to the reaction, or it can be done *in vivo* upon transformation (Li and Elledge, 2007).

### 1.2.3 PCR-based cloning

PCR-based methods are more modular and enable easier assembly of large fragments. One such method is Circular Polymerase Extension Cloning (CPEC), in which the backbone and the insert or inserts to assemble are designed to contain overlapping regions so that during PCR both strands will be denatured and those overlapping regions will hybridise, enabling the DNA polymerase to extend using each other fragment as template. At the end of assembly, there will only be one nick in the sequence, which will be repaired upon transformation.

Figure 1-2 illustrates how the process works. The overlapping regions are typically 25 bp long and have melting temperatures of 60 – 70 °C (Quan and Tian, 2009). CPEC requires a PCR protocol with a high-fidelity DNA polymerase and an extension time proportional to the size of the final assembled product; depending on this, the cloning step may take several hours.
Gibson assembly is a quicker method: despite employing a polymerase, it does not require a PCR reaction setup. Like in CPEC, overlapping backbone and inserts are needed. The reaction mix will contain a high-fidelity DNA polymerase, T5 exonuclease and a DNA ligase. T5 exonuclease digests DNA fragments from their 5’ ends, generating single-stranded overhangs that will anneal with each other so that the polymerase can fill in the gaps and the DNA ligase will seal the nicks. The reaction takes place at 50 °C to inactivate T5 exonuclease rapidly and prevent excessive digestion (Gibson et al., 2009). Short fragments may be chewed back excessively; in fact, fragment length is a limitation in Gibson assembly whereas this is not a drawback in CPEC. However, this protocol is shorter than CPEC, taking from 15 min to 1 hour. Figure 1-3 illustrates how Gibson assembly works.
1.2.4 Recombinatorial cloning

Recombination-based methods other than SLIC are usually applied to rapid genome engineering and screening. Recombineering, or recombination-based genetic engineering, is an *in vivo* method for introduction of one genome modification at a time by means of an exogenously added ssDNA that will recombine with the genome. A classic example is the bacteriophage λ recombination system, known as λ Red, which employs three λ phage proteins: Gam, Exo and Bet. Gam inhibits exonuclease V action, and Exo and Bet mediate recombination. This can be used for gene replacements with simple PCR products (Murphy, 1998). A more advanced method employing λ Red
recombineering is multiplex automation genome engineering (MAGE). MAGE enables multiple modifications of multiple loci in a cell population simultaneously using synthetic degenerate oligonucleotides, which can be used as a method for directed evolution and whole-genome recoding (Wang et al., 2009). This has been used in conjunction with conjugated assembly genome engineering (CAGE), which is a method of large-scale assembly of various modified genomes involving conjugation (Isaacs et al., 2011).

An alternative method for genome modifications using yeast is Tandem Repeats coupled with Endonucleolytic Cleavage (TREC), which is based on two events: cassette insertion via recombination, and excision of that cassette (Noskov et al., 2010). Recombination-based methods are more common for large DNA fragments. The JCVI centre routinely uses several yeast-derived recombination methods to assemble large constructs, up to whole genomes: from in vitro recombination (Gibson et al., 2008), to Recombinase-Mediated Cassette Exchange (RMCE), which works in vivo and employs the Cre-loxP system (Noskov et al., 2015).

1.3 The potential of synthetic biology

A mere century ago, the perspective of being able to use genetic engineering to re-engineer microorganisms in order to make them ‘work for us’ was considered science fiction. Now that the necessary tools have been developed in the form of a wide availability of multiple complex DNA manipulation techniques, it is possible to apply synthetic biology to fields such as biomedicine, bioremediation or manufacturing and find cheaper ways to produce chemicals, biofuels and textiles while harnessing renewable resources; develop more targeted therapies to tackle diseases like cancer; and reduce environmental pollution, amongst others. Some of these potential applications are discussed herein.
1.3.1 Therapeutic applications

Synthetic cell-based therapies would ideally have bacteria as ‘physicians’ that migrate to the affected tissue and are then able to diagnose, make decisions on treatment, express genes for treatment and then self-eliminate from the human body (Claesen and Fischbach, 2015). This remains a wish list, since as of now it is unfeasible to address all of these simultaneously and instead most potential therapies focus on one or two aspects.

For instance, bacteria can be engineered to invade a specific cell type: an *E. coli* strain has been engineered to invade cancer cells by constitutively expressing invasin, a protein produced by *Yersinia pseudotuberculosis* which mediates penetration of bacteria into mammalian cells. At high cell densities or in hypoxia, thanks to heterologous environmental sensors, invasin expression enables invasion of several cancer cell types (Anderson et al., 2006). The Weiss lab developed a HeLa type classifier circuit potentially applicable to cancer therapies that goes beyond just cell type detection: this classifier can make a decision on the state of the cell, discerning between healthy and HeLa cells, and act accordingly by triggering apoptosis only in HeLa through expression of apoptotic protein hBax. Assessment of the high-HeLa and low-HeLa states occurs via evaluation of the expression levels of six different miRNAs based on the HeLa expression profile (Xie et al., 2011). Other therapeutic applications focus solely on treatment decision-making: Chen, Jensen and Smolke (2010) developed an RNA-based control device consisting of a ribozyme switch that triggers cytokine expression and subsequent T-cell proliferation in response to a drug input. In absence of that drug, there is no cytokine expression and T-cells enter apoptosis.

Most proof-of-concepts are developed in *E. coli* since this is the model organism and it is well understood, but it might not be the ideal chassis for *in vivo* applications. Bacteriophages are common gene delivery vectors in gene therapy and can be also used in other therapeutic applications: for example, a T7 bacteriophage can be engineered for
enzymatic degradation of biofilms in bacterial infections, a significant medical problem (Lu and Collins, 2007). In addition, commensal bacteria such as *Bacteroides*, *Bifidobacteria*, and *Lactobacillus* are also good candidates since they are currently used as probiotics and are therefore more likely to gain more public acceptance (Claesen and Fischbach, 2015).

1.3.2 Applications in metabolic engineering

The symbiosis of metabolic engineering and synthetic biology provides an opportunity for cost-effective production of drugs and chemicals by means of controlling branch points in endogenous metabolism or inserting heterologous pathways. Biofuel production from renewable resources has benefited from this symbiosis; for instance, isobutanol can be produced in *E. coli* via a non-fermentative synthetic pathway, diverting amino acid intermediates towards 2-keto acid metabolism leading to alcohol production (Atsumi et al., 2008).

The Keasling lab constructed a *S. cerevisiae* strain producing artemisinic acid, precursor for artemisinin, a valuable anti-malarial drug normally extracted from the plant *Artemisia annua* (Paddon et al., 2013). Amyris currently produces artemisinin via this semi-synthetic pathway. This yeast has also been engineered to produce succinic acid due to its robustness against low pH (Otero et al., 2007).

1.3.3 Applications in bioremediation

Rather than the difficulties posed by regulations, some authors agree that the main obstacle for GMO use in bioremediation is the poor *in situ* performance as compared to well-defined laboratory conditions, in addition to the appearance of predators that will displace them. Wild microbial physiology and ecology are complex, and it is imperative for the microorganism to be stable, adapt and colonise in such a poorly defined, changing environment while preventing horizontal gene transfer (Cases and de Lorenzo, 2005).
Environmental applications include solid waste and water treatment, water desalination, CO$_2$ recapture and groundwater decontamination (Schmidt, 2012), but most research largely focuses on real-time whole-cell biosensors that react to certain compounds in the environment and respond with a signal. An example is a biosensor based on xylR, transcriptional activator for the TOL pathway in *Pseudomonas putida*: when toluene, benzene or xylene are present, they bind to XylR and activate promoter *Pu*. This triggers luciferase (*luc*) expression, resulting in bioluminescence reporting contaminant presence in water or soil (Willardson et al., 1998). This was further engineered to include T7 polymerase enhancing luciferase expression coded by the *luxCDABE* operon (de Las Heras and de Lorenzo, 2012).

### 1.4 Orthogonality and biocontainment in synthetic biology

Synthetic biology aims to deliver safe products and therapies while enforcing biosafety through implementation of biocontainment and risk assessment methods. Despite the fact that laboratory strains are unlikely to grow and persist in the wild, there is a risk of horizontal gene transfer since DNA can remain in the environment even after cell death (Lyon et al., 2010). Closed systems such as manufacturing facilities using GMOs for production are safer, since there is a physical barrier, but the risk of accidental release must be addressed since the behaviour of these systems cannot be predicted accurately. Thus, there is an obvious necessity to implement safety measures so that a GMO will not pollute or invade the environment and can still be controlled: while it is true that no disease or environmental disaster has been linked to GMOs, and instead these have always been caused by natural pathogenic microorganisms, it is a consensus in the scientific community that biosafety measures are needed and the intrinsically related concepts of orthogonality and biocontainment must not be forgotten when discussing synthetic biology (Moe-Behrens et al., 2013; Torres et al., 2016; Wright et al., 2013). The term ‘orthogonality’ derives from the Greek *orthogōnios*, or ‘right-angled’. Orthogonality also refers to two matters being completely divergent and non-interacting. In
mathematics and computer science, the concept means that modifications on a system will not interfere with another system. In biological terms this implies organisms that do not interact with their environment due to their fundamental dissimilarities (de Lorenzo, 2011).

There are several biocontainment approaches to enforce orthogonality, ranging from auxotrophic organisms to synthetic nucleic acids. Other concepts remain purely hypothetical: ‘mirror life’, for instance, a form of life with enantiomeric building blocks that would be completely incapable of interfering with natural forms of life due to their different operating system (Forster and Church, 2007).

1.4.1 Biocontainment via auxotrophy

Auxotrophic organisms are often used in microbiological research. Auxotrophy is based on dependence of an organism on an externally supplied compound, so that cell death is almost certain if escape occurs. This can happen either because that compound suppresses expression of a toxic product or because it compensates for a gene deletion. One of the first synthetic biology E. coli auxotrophs employed a cassette with hok, a gene coding for a lethal protein that is active across multiple species. Absence of tryptophan induced expression of hok, regulated by the trp operon (Molin et al., 1987). A similar system employed a xylS-gef gene switch for bioremediation in which Gef expression (a lethal protein) was repressed in the presence of benzoate. Cells would degrade benzoate, so that benzoate depletion would trigger Gef expression and, ultimately, cell death (Contreras et al., 1991).

Auxotrophy can also be achieved by enforcing dependence on synthetic amino acids. Rovner et al. (2015) used a genomically recoded E. coli with no TAG codons on its genome, incorporated TAGs in up to 22 essential genes via MAGE and introduced a tRNA:aminoacyl-tRNA synthetase pair from Methanocaldococcus jannaschii that can place a synthetic amino acid on UAG.
1.4.2 Biocontainment via induced lethality: kill switches

Kill switches induce lethality upon addition of an inducer and are useful as backups prior to release of the GMO: for instance, once a synthetic organism finished cleaning contaminated waters, cell death could be triggered by adding that inducer leaving naturally present organisms unaffected (Moe-Behrens et al., 2013). One example is addition of IPTG to induce a lac promoter controlling hok expression; however, cells can easily acquire resistance to endogenous hok expression, rendering this kill switch not apt for biocontainment (Bej et al., 1988).

Synthetic genetic counters for cell death can be used to design alternative kill switches. Friedland et al. (2009) engineered synthetic genetic counters in E. coli with 2- and 3-counters for expression of different proteins upon arabinose pulses. If tied to the cell biological cycle, cell death would be programmed after a period of time.

Other common approaches consist of nuclease induction to trigger cell death. To name a few examples, Ahrenholtz, Lorenz and Wackernagel (1994) used the extracellular nuclease of Serratia marcescens under a thermoinducible promoter to degrade 80% of intracellular DNA; Torres et al. (2003) designed a dual gene containment circuit with restriction enzyme EcoRI and colicin E3; and Caliando and Voigt (2015) created a genetic device using CRISPR for either plasmid or genome degradation.

1.4.3 Biocontainment via addiction mechanisms

Addiction, or post-segregational killing mechanisms, attempt to avert the risk of cell-free DNA transmission by using killer proteins and their repressors, encoded in a recombinant plasmid and the host genome respectively. If a host acquired the vector, cell death would be triggered as its genome would not encode the immunity protein. Examples of such mechanisms are the hok/sok system (Bej et al., 1988; Molin et al., 1993), colicins (Torres et al., 2000), the CcdA/CcdB toxin-antitoxin system (Bernard and Couturier, 1992), and restriction-modification systems (Balan and Schenberg, 2005). In order to increase
efficiency of this genetic safeguard, this can be coupled with an additional layer of biocontainment: for example, Torres et al. (2003) employ both endonuclease EcoRI and colicin E3, which have different cellular targets (DNA and RNA respectively).

1.4.4 Biocontainment via xenobiology

Most approaches aim to trigger cell death when it trespasses the boundaries set by the biological designer. However, cell death does not necessarily imply DNA destruction, which can endure harsh conditions and potentially be acquired by existing microbes in the environment: not just by naturally competent bacteria, but also by organisms with induced competency. An example is lighting-competent bacteria, which can be isolated from soil samples (Cérémonie et al., 2004). This highlights the need for more sophisticated biocontainment strategies such as xenobiology.

Xenobiology is a form of biology involving synthetic nucleic acids. These can be created through modifications in the sugar backbone, base pairings, internucleotide linkages or base substitutions. Examples are unnatural base pairs such as isoC and isoG (Johnson et al., 2004) and Ds and Px (Yamashige et al., 2011); size-expanded DNA (xDNA) with larger base pairs than natural DNA (Krueger et al., 2011); and alternative backbones with propagation capabilities, such as hNTPs (Vastmans et al., 2002), CeNTPs (Kempeneers et al., 2005) and tNTPs (Ichida et al., 2005). Additional coding systems involve tetra- and penta-codons (Seligmann, 2015), so that proteins coded by those systems are not be synthesised via the natural translation systems.

If the synthesis of a xeno-organism if reached, it could function as a chassis and be further engineered via by adding multi-compartmental vesicles for further modulation and complex process segregation (Elani et al., 2015). Following the idea of genome transplantation (Lartigue et al., 2007), this genome could be transplanted into the host organism, taking over the cell along with its proteome (Schmidt, 2010).
The creation of an XNA-based organism involves several challenges. In such an organism, XNA would be effectively replacing DNA; chemical synthesis of XNA would be followed by its biological synthesis, as well as XNA-binding proteins responsible for replication, transcription and XNA packaging; ATP, CTP and GTP would possibly need removing from the cell metabolism; and gene flow between DNA and XNA would need to be prevented (Schmidt, 2010). Working towards this, some polymerases have been designed to synthesise XNA from DNA templates, reverse transcribe XNA into DNA, and even accept XNAs, synthesising XNA from a XNA (FANA) template (Pinheiro et al., 2012). In addition, some XNAs have been described to bind very weakly to natural nucleic acids, such as homo DNA (Hunziker et al., 1993) and GNA (Meggers and Zhang, 2010).

1.4.5 DNA-free chassis

Currently, the synthesis and maintenance of an XNA genome or a xeno-organism is currently unreachable, and other biocontainment methods still pose a risk of uncontrolled release. In Europe, legislation is unlikely to be modified to accommodate GMO release, as covered by European directive 2001/18/EC on the deliberate release into the environment of genetically modified organisms (2001). This may be circumvented if a DNA-free chassis was developed, as it would not fit the definition of an organism as "any biological entity capable of replication or of transferring genetic material" as defined by the directive.

Chassi containing no DNA can be obtained from certain E. coli strains and other bacteria. Some of these methods have existed since the 1960 and 70s, although they were originally designed for a purpose other than as a tool to circumvent the risk of uncontrolled release. The first method to be described was the Minicell method in 1967, and then Maxicell in 1971.
1.4.5.1 Minicells

Miniature cells, or minicells, were first reported in 1930 but were only analysed and named in the 1960s (Adler et al., 1967). *E. coli* K12 P678-54 was reported to generate spherical cells from the poles of normal rod-shaped cells during exponential growth, which appeared to be product of an invagination of the bacterial cell wall parallel to normal cell division. Minicells were a tenth of the volume of normal cells and could be separated by sucrose gradient centrifugation. Only traces of DNA were present in minicell preparations, which was attributed to contaminating normal cells; however, they contained RNA and proteins. Isolated minicells could not be induced to produced β-galactosidase, but if a mixed culture was induced, minicells containing β-galactosidase could be isolated.

All minicell-producing *E. coli* strains are derived from P678-54. The genetic basis is a Min- phenotype caused by at least a mutation in the *minB* locus (Davie et al., 1984), which was later on reported to be responsible for proteins MinC, MinD and MinE. This Min system has since been extensively characterised and is responsible for placing the division septum at the middle of the cell so that the two daughter cells are the same size (de Boer et al., 1989).

The division septum, or Z ring, is composed of the FtsZ protein and is inhibited by the Min system from either cell pole so that division occurs precisely at the centre of the cell (Bi and Lutkenhaus, 1990). Mutations in the Min system prevent formation of the division septum at midcell, instead occurring at the polar ends of normal cells, leading to one chromosome-containing daughter cell and another chromosome-less daughter cell, the minicell. This is illustrated in Figure 1-4. Overproduction of FtsZ alone has been described to also induce minicell formation (Ward and Lutkenhaus, 1985). If there is plasmid DNA at the point of invagination, minicells may retain this. Jivrajani, Shrivastava and Nivsarkar (2013) took advantage of this and included a GFP-coding plasmid in
E. coli minicell-producing strain PB114 to facilitate imaging of the minicell population and evaluate their purification protocol.

Inactivation of the Min system in other species enables production of minicells. In fact, minicell production has been reported in other Gram-negative bacteria such as *Shigella flexneri*, *Salmonella enterica* and *Pseudomonas aeruginosa* (MacDiarmid et al., 2007); and Gram-positives, like *Listeria monocytogenes* (MacDiarmid et al., 2007), *Bacillus subtilis* (Reeve et al., 1973) and *Corynebacterium glutamicum* (Lee et al., 2015).

![Figure 1-4. Illustration of minicell formation through abnormal cell division. The Z ring is not formed at the middle of the cell, generating a normal cell and a smaller, chromosome-less cell that may retain plasmid DNA. Adapted from (Carleton et al., 2013).](image)

Initially, minicell purification protocols included a sucrose gradient centrifugation step (Adler et al., 1967; Cohen and Cohen, 1968), which is laborious, does not guarantee a solution free of parental cells and has a low yield. The addition of penicillin in minicell protocols (Levy, 1970) improved yields by inducing lysis of cells with a growing cell wall. Protocols used currently incorporate multiple centrifugation steps, filtration and use of penicillin-derived antibiotics. A more efficient approach incorporates ceftriaxone, a cephalosporin affecting cells with a growing cell wall but no minicells, and filtration to obtain purified minicells free of contaminating parental cells with a high yield (Jivrajani et al., 2013). Other approaches include growing *E. coli* in stressful conditions like high salt conditions.
concentration in order to promote filamentation, increasing the size difference between regular cells and minicells and facilitating purification (MacDiarmid et al., 2007).

Minicells have been proposed as chassis for biosensors: since minicells lack chromosomal DNA but retain plasmid DNA and transcription and translation machinery they can respond to signals in the environment. Minicells with synthetic gene circuits have been shown to express GFP in response to inducers like L-arabinose, acrylate or glucarate (Rampley et al., 2017) and aspirin (Chen et al., 2019). *E. coli* minicells have also been used as recipients for conjugation: Cohen and Cohen (1968) showed an increase in dsDNA with respect of ssDNA post-conjugation, and more recently Kobayashi (2018) regenerated minicells to viable *E. coli* and confirmed their genome to be that of the donor via pulse-field gel electrophoresis (PFGE).

1.4.5.2 Maxicells

The Maxicell method, first published in 1979 (Sancar et al., 1979), was used extensively in the 80s and early 90s for identification of proteins coded by specific DNA sequences inserted into plasmids. In this method, *E. coli* was irradiated with UV light to produce extensive damage to the chromosome preferentially, leaving plasmid DNA mostly unaffected. Addition of a radiolabelled amino acid allowed identification of proteins coded by the plasmid since the chromosome would be too damaged to permit transcription of chromosomal genes. For this purpose, *E. coli* strains with impaired DNA repair were used. The most common strain was *E. coli* CSR603, known as the Maxicell strain, since its *recA1 uvrA6 phr-1* genotype means it has weakened nucleotide excision repair, recombination repair and photoreactivation.

Proteins identified using maxicells include DNA repair protein (UvrA) (Kacinski et al., 1981), *E. coli* membrane protein diacylglycerol kinase alpha (DgkA) (Loomis et al., 1985), lipoyl synthase (LipA) (Vanden Boom et al., 1991), and \( \alpha \) ketoglutarate-dependent dioxygenase (AlkB) (Kataoka and Sekiguchi, 1985). The Maxicell method is no longer used due to availability of high-throughput sequencing technologies and computational
tools. In addition, methods employing UV light are not desirable due to limited scalability and reproducibility.

Methods other than UV irradiation can be used to produce maxicells. Microcin B17, a DNA gyrase poison, can be added instead to induce DNA damage (Mayo et al., 1988). A variant, the chemically induced Maxicell method, uses the capacity of rifampicin to inhibit translation of *E. coli* RNA polymerase-dependent genes and the resistance of T7 polymerase to this antibiotic. A gene under control of T7 promoter is introduced in a plasmid, and T7 RNA polymerase is present either in the *E. coli* genome as a T7 prophage or under control of IPTG induction in a different plasmid. When rifampicin is added, only T7 RNA polymerase will be active and therefore only those genes under control of T7 promoter will be transcribed (Studier and Moffatt, 1986).

Maxicells can become conjugation donors by retrotransfer. In order to do this, maxicells are prepared from CSR603 cells carrying a mobilizable non-self-transferable plasmid. A donor with a self-transferable plasmid is mated with these maxicells, which receive this plasmid and become retrodonors that are able to transfer the mobilizable plasmid to the original donors, now retrorecipients (Heinemann and Ankenbauer, 1993).

### 1.4.6 Resistance to biocontainment measures

Cells can overcome genetic firewalls in different manners. Auxotrophs may be perfectly contained in a regulated laboratory environment, but they can survive on secondary metabolites of other organisms and small molecules present in the environment (Ritger et al., 2011). Genetic switches can be overcome if mutations arise in the population, and leaked expression of the lethal gene can exert selective pressure over the population to develop immunity against it (Knudsent and Karlstrom, 1991). Incomplete induction may also occur and uninduced cells can persist in the population: Ahrenholtz, Lorenz and Wackernagel (1994) designed a thermoinducible kill switch and obtained several
surviving clones that did not survive a second induction, hypothesising that they were not successfully induced the first time.

Not all biocontainment measures are infallible. Moe-Behrens, Davis and Haynes (2013) compiled the escape frequencies of engineered auxotrophs and bacteria with induced lethality switches and compared them to the guidelines recommended by the NIH, in 2 L (National Institutes of Health, 2016). Out of the 13 works, 2 did not report an escape frequency; 2 were under that maximum; 2 just in the boundary; and 7 were above it. This highlights the need for further work and optimisation of current genetic safeguards.

1.5 DNA damage in bacteria

Among the genetic safeguards discussed in section 1.4, kill switches that depend on DNA degradation are the most relevant to this work. Several DNA-damaging genetic circuits have been created to trigger DNA degradation and subsequent cell death, as discussed in section 1.4.2. However, activation of these circuits does not guarantee complete elimination of DNA and horizontal gene transfer must be prevented in all cases, which does not come accompanied by cell death. Therefore, a kill switch that triggered complete DNA degradation and cell death would enable creation of a chassis that could be released with no fear of interference with the environment. In order to design this DNA-free chassis, it is necessary to first understand how DNA damage occurs and how bacteria attempt to repair it.

DNA damage can have two different sources: it can be part of normal cellular processes, due to reactive oxygen species or replication errors; or caused by external agents, such as the ones discussed in section 1.5.2. Damaged DNA may lead to lethal mutations and cell death and all species have evolved to have repair mechanisms, which are similar across species and even in both prokaryotes and eukaryotes. This section will be focusing on double (DSBs) rather than single-strand breaks (SSBs) since they are more relevant to this study.
1.5.1 Double-strand breaks (DSBs)

Double-strand breaks (DSB) are incisions in a DNA duplex which can normally be repaired through different mechanisms. If repaired inefficiently, a DSB can cause mutations or loss of genetic information. Cell death may occur upon a high incidence of DSB due to an impossibility to perform DNA replication, transcription and protein synthesis.

DSB induction has been exploited as a form of genetic firewall, i.e. as part of an addiction mechanism (Caliando and Voigt, 2015; Torres et al., 2003); to study DNA repair pathways, such as non-homologous end joining (NHEJ); or as part of the Maxicell method (Sancar et al., 1979). A range of DSB-inducing agents may be optimised for rapid cell death or as a mechanism for DNA ablation. These agents are summarised in Table 1-1 and are discussed herein.

Table 1-1. Summary of the main sources of DSBs in bacteria.

<table>
<thead>
<tr>
<th>Agent</th>
<th>Examples</th>
<th>Action on DSB</th>
<th>Sequence specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type-II REs</td>
<td>Smal, EcoRI, BamHI</td>
<td>Direct cleavage when sequence is unmethylated</td>
<td>Yes; 4-8 bp</td>
</tr>
<tr>
<td>Homing endonucleases</td>
<td>I-Drel</td>
<td>Direct cleavage upon recognition</td>
<td>Yes; 12-40 bp</td>
</tr>
<tr>
<td>CRISPR-Cas</td>
<td></td>
<td>Direct cleavage when complex binds to target sequence</td>
<td>Yes; ~40 bp</td>
</tr>
<tr>
<td>ROS</td>
<td></td>
<td>Indirect: mostly due to replication fork stall</td>
<td>No</td>
</tr>
<tr>
<td>UV radiation</td>
<td>DNAse colicins (E2, E7, E8, E9)</td>
<td>Unrepaired gaps skipped by polymerases</td>
<td>No</td>
</tr>
<tr>
<td>Bacteriocins</td>
<td></td>
<td>Cleavage when unbound to cognate immunity protein</td>
<td>No</td>
</tr>
<tr>
<td>DNA gyrase inhibitors</td>
<td>Nalidixic acid, ciprofloxacin (subunit A)</td>
<td>Replication fork arrest by either enzymatic inhibition or DNA-enzyme complex stabilisation</td>
<td>No</td>
</tr>
<tr>
<td>Radiomimetic antibiotics</td>
<td>Bleomycin, neocarzinostatin</td>
<td>Indirectly, following SSBs in both strands</td>
<td>Variable, eg GT dinucleotides for bleomycin</td>
</tr>
<tr>
<td>Alkylating agents</td>
<td>Mitomycin C</td>
<td>Indirectly, following SSBs in both strands</td>
<td>No</td>
</tr>
</tbody>
</table>
1.5.2 DNA damage inducing agents

1.5.2.1 Endonucleases

Endonucleases are enzymes that catalyse the cleavage of the phosphodiester bond of a polynucleotide chain in a specific or non-specific manner. Bacteriophages encode endonucleases in order to degrade the host genome and scavenge nucleotides for its own genome replication; an example is bacteriophage T4 endonuclease II, denA (Hercules et al., 1971).

Restriction enzymes are a type of specific endonucleases and comprise a system for cell defence against foreign agents by means of exogenous DNA restriction upon recognition of 4, 6 or 8 bp targets. A restriction-modification gene complex is composed of a restriction endonuclease and its cognate modification enzyme. This modification enzyme is a methyltransferase acting on a specific target sequence, which once methylated is protected from restriction. Exogenous DNA will not be methylated with that specific pattern, being susceptible of restriction. Restriction enzymes used in molecular biology are mostly type-II, in which methyltransferase and endonuclease are separate activities and they cleave close to the recognition sequence. In addition, they have been used as a conditional suicide system (Torres et al., 2003).

Homing endonucleases have larger targets than restriction enzymes (12 - 40 bp) and are coded in introns or inteins. They are flexible in terms of sequence recognition, so they can recognise sequence families. In therapeutic applications, they are considered safer than restriction enzymes due to their longer sequence specificity, which reduces the risk of off-targeting. Homing endonucleases can be designed to target a defective gene in a monogenic disease, cleaving the sequence and triggering repair through homologous recombination (Marcaida et al., 2010).

Pulse-field gel electrophoresis (PFGE) is a DNA electrophoresis technique that enables separation of large fragments by periodically switching the direction of the voltage. For
this reason, it can be used to combine enzymatic restriction of a whole genome and high-resolution separation of the resulting bands. Large fragments are produced by using a rarely-cutting restriction enzyme (Löbrich, M., Ikpeme, S., & Kiefer, 1994). Such enzymes have 8 bp targets, or recognise targets containing only A and T or C and G: examples are Ascl, NotI, SapI, SgrAI or Swal (Hanak, Alexis & Ward, 2001). Thus, PFGE is known for its application on DNA fingerprinting and study of genetic diversity within a population (Otlewska et al., 2013; Zou et al., 2013). Additionally, this technique was used prior to availability of genome sequencing for genome size estimation, for example in frequent P. aeruginosa strains (Hector and Johnson, 1990). Field-inversion gel electrophoresis (FIGE) is a type of PFGE in which the direction of the electric field is periodically switched at a 180° angle.

1.5.2.2 CRISPR-Cas
CRISPR-Cas is an adaptive bacterial immune system that functions similarly to RNA interference. This system employs RNA-guided nucleases (Cas) to cleave invading DNA. In a natural CRISPR-Cas system, when foreign DNA enters the cell, crRNA is produced which contains the protospacer regions. TracrRNA, also produced in this system, forms a duplex with crRNA and then recruits Cas9 nuclease. This complex then cleaves the sequence complementary to crRNA. Engineered variations of this system employ a fusion of crRNA and tracrRNA, forming gRNA (guide RNA) which will then recruit Cas9. The recognised sequence must be next to a protospacer adjacent motif sequence (PAM) (Jinek et al., 2012). CRISPR-Cas systems recognise a longer sequence than restriction enzymes, approximately 20 bp, of which 12 bp at the 3’ end are crucial for cleavage specificity (Jiang et al., 2013).

1.5.2.3 Reactive oxygen species
Despite SSBs are the predominant DNA breakage induced by reactive oxygen species (ROS), DSBs are the main cause of cell death. NHEJ-deficient cells show substantial chromosomal DSB unrepaired damage as a result of oxygen metabolism (Karanjawala
et al., 2002). DSB are unlikely to be a direct consequence of ROS damage (Jeggo and Löbrich, 2007); rather, DSB occur during replication in areas with previous ROS-induced lesions or when this lesion is encountered by transcription machinery leading to stalled forks (Woodbine et al., 2011).

1.5.2.4 UV damage
UV light cause diverse lesions in DNA, including cyclobutane-pyrimidine dimers (CPDs) and pyrimidine adducts called 6-4 photoproducts (6-4PPs), the former being more abundant, as well as their Dewar valence isomers (Sinha and Häder, 2002). These lesions affect DNA replication and transcription, causing polymerases to skip affected regions and leading to gaps in the newly replicated strand. If unrepaired, they will result in a DSB (Smith, 2013). As mentioned in section 1.4.5.2, UV-induced DNA photodamage was the underlying mechanism behind the Maxicell method (Sancar et al., 1979).

1.5.2.5 Bacteriocins
Bacteriocins are a large and diverse group of bacterial peptides and proteins targeting closely related strains. These range from small peptides like lanthipeptides produced by lactic acid bacteria, to larger oligomeric proteins such as colicins. Their narrow spectrum makes them safe for humans: bacteriocin nisin produced by Lactococcus lactis received GRAS (generally recognised as safe) status in 1988 for food applications (FDA, 1988).

Colicins are modular bacteriocins produced by E. coli which have three functional domains for binding to the outer membrane, crossing and translocating to their site of action. The site of action is dependent on their specific function: enzymatic colicins degrade cytoplasmic host nucleases; pore-forming colicins depolarise the cytoplasmic membrane; and colicin M inhibits peptidoglycan synthesis in the periplasm (Kim et al., 2014). Class E colicins bind to vitamin B12 receptor BtuB, and include enzymatic colicins with DNase (E2, E7, E8, E9) and RNase (E3, E4, E6) activities (de Zamaroczy and Buckingham, 2002).
Colicins are usually coded in plasmids together with an Immunity protein (Imm), which inhibits the colicin when bound to it. Colicins have been used in addiction mechanisms as a second layer of containment (Torres et al., 2003) and could be used as a kill switch if expressed without their immunity protein.

### 1.5.2.6 DNA gyrase inhibitors

DNA gyrase is an essential protein in replication and transcription. Its function is to induce negative supercoilings in DNA by means of transiently inducing a DSB in presence of ATP, necessary to pass one strand through the other. Subunit GyrA is involved in DNA cleavage and strand passage, and subunit GyrB is involved in ATP binding (Cozzarelli, 1980). DNA gyrase inhibition leads to replication fork arrest, which can result in an unresolved DSB. DNA gyrase inhibitors can act either by inhibiting the enzymatic action of gyrase or by stabilising the covalent DNA-enzyme complex.

Different inhibitors will bind to either subunit A or B, and *E. coli* resistant mutants for both subunits have been identified: for example, nalidixic acid and ciprofloxacin bind to GyrA and stabilise the DNA-enzyme complex, while novobiocin targets GyrB preventing DNA. Other proteins targeting GyrB include Ccdb, part of a toxin/antitoxin system, and microcin B17 (MccB17). MccB17 is a small post-translationally modified peptide that stabilises the covalent enzyme-DNA complex, leading to cleavage.

### 1.5.2.7 Radiomimetic antibiotics

Radiomimetic antibiotics act through deoxyribose oxidation in both DNA strands. Bleomycin is a well-known radiomimetic drug produced by fermentation of *Streptomyces verticillus* and used in cancer treatments such as Hodgkin’s lymphoma. Bleomycin binds to DNA adjacent to guanosines, interfering with replication and transcription and causing SSBs and DSBs (Packer, 1994). Despite DSBs are considered a secondary effect of bleomycin, as there is a requirement of a SSB in the primary site for attack at the complementary strand following bleomycin reactivation, they are the primary cause for cytotoxicity (Steighner and Povirk, 1990).
Another radiomimetic antibiotic is neocarzinostatin, which causes SSBs and DSBs by intercalation between base pairs and subsequent hydrogen abstraction. In this case, DSBs may occur either directly or, most commonly, via an abasic site with a closely opposed SSB (Povirk, 1996).

1.5.2.8 Alkylating agents
Alkylating agents such as mitomycin C (MMC) act by adding an alkyl group to the N7 of guanine in DNA, destabilising the glycosidic bond and resulting in an AP site. If the ribose ring opens, it results in an SSB. If this happens in the same area of both DNA strands, a DSB is formed (Roberts et al., 2014). MMC can also induce cross-linking in CpG-rich sites (Teng et al., 1989).

1.5.3 DNA repair in prokaryotes
Since unrepaired DNA damage can lead to cell death, cells have naturally developed strategies for DNA repair. These pathways are highly conserved in prokaryotes and eukaryotes, but this section will focus on bacterial DNA repair pathways. Cell death can ultimately happen if the amount of DNA damage is overwhelming to the repair mechanisms of the cell and thus strains with altered repair pathways are more sensitive to DNA damaging agents.

1.5.3.1 Photoreactivation
The photoreactivation pathway is absent in humans and occurs only in presence of light. It involves photolyases repairing UV-induced cyclobutane pyrimidine dimers by separating them, cleaving the bond between two pyrimidines. The mechanism is based on binding and dimer cleaving with the energy gained from light, hence restoring the bases to their original form (Snyder and Champness, 2013).

1.5.3.2 Base excision repair (BER)
The BER pathway repairs damage caused by ROS, hydrolysis and other metabolites. DNA glycosylases remove the damaged base, and the AP (apurinic/apyrimidinic) site is
removed by either an AP lyase or endonuclease, which nicks the DNA strand 3’ or 5’ to the AP site respectively. The gap is filled by a DNA polymerase and sealed with a DNA ligase (Seeberg et al., 1995). DSBs occur mainly in clustered DNA lesion sites whilst the cell attempts to repair them.

**1.5.3.3 Nucleotide excision repair (NER)**

The NER machinery recognises DNA helix distortions rather than alterations in its chemical composition, which together with its low specificity renders it a very important pathway effective against many types of DNA damage. This pathway is present in most organisms, highly conserved and much more complex in eukaryotes as compared to prokaryotes, involving up to 30 genes (Sinha and Häder, 2002).

In bacteria, the proteins involved are UvrA, UvrB and UvrC in a complex called the UvrABC endonuclease. Two copies of UvrA dimerise and bind one UvrB, forming a complex that binds non-specifically to DNA until it encounters a helix distortion; UvrA is then replaced by UvrC, nicking DNA 7 nucleotides upstream and 4 nucleotides downstream of the damage. Helicase UvrD removes the resultant oligonucleotide. DNA polymerase I fills in the gap, and the duplex is restored by ligase (Van Houten, 1990).

**1.5.3.4 SOS response**

The SOS response is a last-resort mechanism after intensive DNA damage in which the cell attempts to generate a mutation that will enable survival. It is triggered by the action of a series of proteins: in normal conditions, the LexA repressor is tightly bound to the SOS box, sequences upstream of SOS genes promoters. When it detects ssDNA, RecA nucleofilaments this and activates, promoting the cleavage of LexA, freeing the SOS box and enabling transcription of SOS genes. There are over 40 SOS genes, including low-fidelity polymerases that replicate damaged DNA introducing many mutations (Napolitano et al., 2001).
1.5.3.5  **Recombination repair**

Homologous recombination is a process in which two similar sequences are exchanged in either single or double strands. In bacteria, the RecBCD pathway repairs DSBs, and the RecF pathway repairs SSBs.

The RecBCD complex, also called exonuclease V, has both helicase and nuclease functions. When it encounters a free DNA end, its helicase activity unwinds the DNA duplex and then exerts its ATP-dependent 3’-5’ exonuclease activity until it encounters and passes an 8-nucleotide Chi site or recombinational hotspot. RecBCD then stops degrading DNA, and RecA binds nucleating the ssDNA regions. This RecA-ssDNA filament invades a homologous dsDNA, initiating recombination (Anderson and Kowalczykowski, 1997). In the RecF pathway, the RecFOR complex acts similarly to the RecBCD complex and might compensate its deficiency to some extent (Hiom, 2009).
1.6 Aims and objectives

Despite the number of biocontainment strategies reported in the literature, not one completely nullifies the risk of unintended transfer of genetic material or uncontrolled replication through spontaneous mutation and unpredictable behaviour outside a controlled laboratory environment. This thesis proposes a novel approach to biocontainment: genome ablation, a strategy for degradation of DNA in *E. coli* with preservation of metabolic activity.

The specific aims of this work are as follows:

1- To develop a strategy for *in vivo* DNA degradation. This requires exploration of DNA-degrading proteins, exonucleases, in combination with other DNA-cleaving agents such as endonucleases, antibiotics, and UV light. A reliable method of DNA quantification will be employed to measure the extent of degradation.

2- To prove that expression of DNA-degrading proteins or addition of DNA-damaging agents results in a decrease in cell viability, and that implementation of the genome ablation strategy results in a chassis unable to replicate.

3- To explore the applications of this genome-ablated chassis through constitutive expression of proteins of easily detectable activity and evaluation of protein activity post-genome ablation, demonstrating applicability of this strategy to fields such as industrial biocatalysis or bioremediation.

4- To evaluate this chassis as recipient for bacterial transformation, proving its ability to express a protein encoded within that plasmid post-genome ablation.

5- To highlight the challenges and limitations of toxic protein production and explore the strategies that cells employ to circumvent it.
## 2 Chapter 2. Materials and Methods

### 2.1 Materials

#### 2.1.1 Strains

All bacterial strains used in the present work are summarised in Table 2-1.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Origin</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli CSR603</strong></td>
<td>F- λ- thr-1 araC14 leuN6(Am) Δ(gpt-proA)63 lacY1 tsx-33 glnX44(AS) phr-1 galK2(Oc) Rac-0 gyrA98(NalR) recA1 rpsL31(Str6) kgdK51 xylA5 mtl-1 argE3(Oc) thiE1 uvrA6</td>
<td>Used for genome ablation experiments due to its impaired DNA repair.</td>
<td>Purchased from the Leibniz Institute DSMZ (Germany).</td>
</tr>
<tr>
<td><strong>E. coli DH5αZ1</strong></td>
<td>lacIqPN25-tetR Sp6 deoR supE44 Δ(lacZYA-argFV169) Phi80 lacZΔM15 hsdR17(RK- mK+) recA1 endA1 gyrA96 thi-1 relA1</td>
<td>Used for cloning and when employing the dual recombinase plasmid system.</td>
<td>Gift from Dr Thomas Folliard (ExcepGen)</td>
</tr>
<tr>
<td><strong>E. coli HB101</strong></td>
<td>F- thi-1 hsdS20 (rB-, mB-) supE44 recA13 ara-14 leuB6 proA2 lacY1 galK2 rpsL20 (Str6) xyl-5 mtl-1</td>
<td>Used when working with methyltransferases for cloning and expression.</td>
<td>Available in the Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td><strong>E. coli TOP10</strong></td>
<td>F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu)7697 galU galK λ- rpsL(Str6) endA1 nupG</td>
<td>Used for cloning.</td>
<td>Purchased from Thermo Fisher Scientific (Basingstoke, UK).</td>
</tr>
</tbody>
</table>

*Table 2-1. Strains used in this work, together with their genotype (when known; the genotype in bold is relevant to the strain and its use), origin and explanation on how they were used.*
<table>
<thead>
<tr>
<th>Organism</th>
<th>Source of template DNA for amplification of the restriction enzyme and methyltransferase.</th>
<th>Available in the Ward lab, Department of Biochemical Engineering, University College London.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>E. coli MG1655</strong></td>
<td>Source of template DNA for exonuclease III gene amplification.</td>
<td>Available in the Ward lab, Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td></td>
<td>F- λ–  ilvG- rfb-50 rph-1</td>
<td></td>
</tr>
<tr>
<td><strong>Nocardia otiditidis-caviarum</strong></td>
<td>Source of template DNA for amplification of the NotI restriction enzyme and methyltransferase.</td>
<td>Available in the Ward lab, Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td></td>
<td><em>Strain and genotype unknown</em></td>
<td></td>
</tr>
<tr>
<td><strong>Proteus vulgaris</strong></td>
<td>Source of plasmid pPVU1, template DNA for amplification of the PvuII restriction enzyme and methyltransferase.</td>
<td>Available in the Ward lab, Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td></td>
<td><em>Strain and genotype unknown</em></td>
<td></td>
</tr>
<tr>
<td><strong>Bacillus amyloliquefaciens</strong></td>
<td>Source of template DNA for amplification of the BamHI restriction enzyme and methyltransferase.</td>
<td>Available in the Ward lab, Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td></td>
<td><em>Strain and genotype unknown</em></td>
<td></td>
</tr>
<tr>
<td><strong>Haemophilus influenzae</strong></td>
<td>Source of template DNA for amplification of the HpaII restriction enzyme and methyltransferase.</td>
<td>Available in the Ward lab, Department of Biochemical Engineering, University College London.</td>
</tr>
<tr>
<td></td>
<td><em>Strain and genotype unknown</em></td>
<td></td>
</tr>
</tbody>
</table>
2.1.2 Plasmid DNA

All plasmids used as template or backbone are summarised in Table 2-2.

Table 2-2. Plasmids used as template DNA, cloning or expression vectors in this work with their relevant characteristics and source. Abbreviations: Amp (ampicillin), Cam (chloramphenicol), Km (kanamycin), Tet (tetracycline), ori (origin of replication).

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristics</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pUC19</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;; lac promoter; multiple cloning site (MCS); lacZα; ColE1 ori</td>
<td>Norrander et al., 1983</td>
</tr>
<tr>
<td>pTTQ18</td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;; tac promoter; MCS; lacZα; lacI&lt;sup&gt;Iq&lt;/sup&gt;; ColE1 ori</td>
<td>Stark, 1987</td>
</tr>
<tr>
<td>pQR445</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;; tac promoter; MCS; lacI&lt;sup&gt;Iq&lt;/sup&gt;; RSF1010 ori</td>
<td>Irvine, 2005</td>
</tr>
<tr>
<td>pQR801</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;; T7 promoter; Chromobacterium violaceum 2025 TAm gene with a N-terminal His-tag; lacI&lt;sup&gt;Iq&lt;/sup&gt;; pBR322 ori</td>
<td>Kaulmann et al., 2007</td>
</tr>
<tr>
<td>pQR226</td>
<td>Km&lt;sup&gt;R&lt;/sup&gt;; Pseudomonas meta pathway; pUC ori</td>
<td>Jackson, 1996</td>
</tr>
<tr>
<td>pACYC184</td>
<td>Tet&lt;sup&gt;R&lt;/sup&gt;; Cm&lt;sup&gt;R&lt;/sup&gt;; p15A ori</td>
<td>Chang and Cohen, 1978</td>
</tr>
<tr>
<td><strong>Dual recombinase system - controller plasmid</strong></td>
<td>Cm&lt;sup&gt;R&lt;/sup&gt;; araB promoter, TP901 integrase; tetR PLTetO1 promoter, Bxb1 integrase; colE1 ori</td>
<td>Dr Thomas Folliard (ExcepGen)</td>
</tr>
<tr>
<td><strong>Dual recombinase system - expression plasmid</strong></td>
<td>Amp&lt;sup&gt;R&lt;/sup&gt;; Bxb1 atts, sfGFP; mkate2, TP901 atts; pSC101 ori</td>
<td>Dr Thomas Folliard (ExcepGen)</td>
</tr>
<tr>
<td>pPVU1</td>
<td>PvuII restriction-modification system; pSC101-like ori</td>
<td>-</td>
</tr>
</tbody>
</table>

2.1.3 Chemicals

All chemicals and antibiotics were purchased from Sigma-Aldrich (Gillingham, UK) unless otherwise stated.

2.1.4 Oligonucleotides

All oligonucleotides were purchased from either Eurofins Genomics (Ebersperg, Germany) or Sigma-Aldrich (Gillingham, UK). A list of the sequences of all oligonucleotides used in this work can be found in the Appendix, section 8.2.
2.1.5 gBlocks

gBlocks were supplied by Integrated DNA Technologies (IDT, Leuven, Belgium). Tubes containing the gBlocks were centrifuged for a minute at 13,800 x g and resuspended to 10 ng / µL in Tris-EDTA buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0). A list of the sequences of all gBlocks used in this work can be found in the Appendix, section 8.4.

2.1.6 Antibiotic stocks

All antibiotics stocks were dissolved in either molecular biology-grade water or 99 % ethanol, filtered through a 0.2 µm filter and stored at – 20 °C. The stock concentrations were as follows: ampicillin (Amp) 100 mg / mL; chloramphenicol (Cam) 34 mg / mL; kanamycin (Km) 50 mg / mL; tetracycline (Tet) 10 mg / mL. All antibiotics were used at a 1:1000 dilution.

2.1.7 Growth media

Luria-Bertani (LB) broth was prepared at 25 g / L in ultrapure water and then autoclaved. In order to prepare LB agar plates, agar was also added at 15 g / L and autoclaved.

M9 media was prepared with M9 salts (6.8 g / L disodium phosphate, 3 g / L monopotassium phosphate, 1 g / L ammonium chloride, 0.5 g / L sodium chloride), 1 % casamino acids, 1 mM magnesium sulphate, 0.1 mM calcium chloride, and 1 % D-glucose in ultrapure water. All components were filter-sterilised prior to combining. M9 agar was prepared with a 2X M9 media solution mixed with a warm autoclaved 30 g / L agar solution.

In order to prepare both LB and M9 plates, the media was cooled to 50 ºC before adding the antibiotics and pouring plates.
2.2 Methods

2.2.1 DNA manipulation and analysis

2.2.1.1 Sequencing
All DNA samples were diluted to a concentration of 80 – 90 ng / µL and sequenced via the Eurofins Genomics TubSeq service (Ebersberg, Germany) with the oligonucleotides specified in the Appendix, section 8.2.1.

2.2.1.2 DNA preparation for cloning
Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (QIAGen, Manchester, UK). Gel extraction of DNA fragments from agarose gels was performed using the Monarch DNA Gel Extraction Kit (NEB, Hitchin, UK). PCR purification and desalting were performed using the DNA Clean & Concentrator Kit (Zymo Research, Freiburg, Germany). If needed, digested DNA was dephosphorylated by addition of calf intestinal alkaline phosphatase, CIP (NEB, Hitchin, UK) at a ratio of 1 unit per pmol of DNA ends and a 30 min incubation at 37 °C.

2.2.1.3 Restriction enzyme digestions
All restriction enzymes were purchased from NEB (Hitchin, UK). Digestion reactions were prepared with the buffer recommended by NEB and with at least 10 units of restriction enzyme per µg of DNA, then incubated at the recommended temperature from 2 h to overnight. Digestions were either inactivated at 80 °C for 20 min, purified with the DNA Clean and Concentrator kit (Zymo Research), or run in an agarose gel.

2.2.1.4 Agarose gel electrophoresis
Agarose gels were prepared in 1X Tris-borate-EDTA buffer (1 M Tris, 0.69 M boric acid, 0.01 M EDTA) at 1 % agarose, or 1.5 % agarose for fragments in the 100 – 200 bp range. If pre-staining, ethidium bromide at 0.2 µg / mL or GelRed at 1X (Biotium, California, US) were added prior to pouring the agarose mixture into the gel casting tray. DNA samples were combined with loading dye (NEB, Hitchin, UK) and then loaded into the wells. Gels were run for either 1 h at 100 V or 1.5 h at 90 V and visualised under UV
light at 365 nm with an Alphalmager Mini (ProteinSimple, Oxford, UK). If post-staining, gels were stained in a 3X GelRed solution at 0.1 M sodium chloride for 30 min prior to visualisation.

2.2.1.5 Densitometric quantification of samples in agarose gels

Agarose gel images were analysed with the software ImageJ (Schneider et al., 2012). Images were converted to black and white and lanes were delimited by a rectangle around the intact bands, discarding the smeared DNA. Lanes were plotted, and the peaks corresponding to DNA present in the gel were selected and quantified.

2.2.1.6 PCR

PCR reactions were set up with 1X Q5 Hot-Start Master Mix (NEB, Hitchin, UK), forward and reverse oligonucleotides at 10 µM, 10-100 ng template DNA, and molecular biology-grade water up to the desired volume. The thermocycling reaction was conducted in a BioRad C1000 Touch Thermal Cycler (Bio-Rad, Watford, UK) as follows: initial denaturation for 30 s at 98 °C; 35 cycles of 10 s at 98 °C, 30 s at the annealing temperature, and 30 s per kb amplified at 72 °C; and a final extension of 5 min at 72 °C. Amplicon size was verified by running 1 to 10 µL of the reaction in an agarose gel as explained in section 2.2.1.4. Conditions for each PCR can be found in the Appendix, section 8.3. A list of the sequences of all oligonucleotides used in this work can be found in the Appendix, section 8.2.

2.2.1.7 Colony PCR

A portion of a colony if bigger than 1 mm, or a whole colony if smaller than 1 mm, was picked with a 200 µL tip or a 1 µL inoculation loop and resuspended in 10 µL molecular-grade water. 1 µL of this was transferred to a PCR tube with 5 µL Taq Hot-Start Master Mix 2X (NEB, Hitchin, UK), 0.25 µL primer stock at 10 µM, and 3.75 µL molecular-grade water. The thermocycling reaction was conducted in a BioRad C1000 Touch Thermal Cycler (Bio-Rad, Watford, UK) as follows: colony lysis and initial denaturation for 5 min 30 s at 95 °C; 30 cycles of 30 s at 95 °C, 1 min at the annealing temperature, and 1 min
per kb amplified at 68 °C; and a final extension of 5 min at 72 °C. Conditions for each PCR can be found in the Appendix, section 8.3.

2.2.1.8  **Ligation**
Ligations were prepared by combining 50 ng backbone and insert at a 9-fold molar ratio in a 20 µL reaction with 400 units T4 ligase in 1X T4 ligase buffer (NEB, Hitchin, UK). The ligations were either incubated at 16 °C overnight, or in a cycled ligation protocol (150 cycles of 3 min at 15 °C and 3 min at 37 °C). In both cases an inactivation step at 80 °C for 20 min was added.

2.2.1.9  **Circular Polymerase Extension Cloning (CPEC)**
CPEC reactions were prepared by combining 50 ng backbone and insert at a 3-fold molar ratio in a 20 µL reaction at 1X Q5 Hot-Start master mix (NEB, Hitchin, UK). CPEC was conducted in a BioRad C1000 Touch Thermal Cycler (Bio-Rad, Watford, UK) as follows: initial denaturation for 30 s at 98 °C; 5 cycles of 10 s at 98 °C, 30 s at 55 °C, and 30 s per kb of assembled vector at 72 °C; and a final extension of 10 min at 72 °C. 5 µL of the CPEC reaction was transformed via chemical transformation (see section 2.2.4) or desalted via drop dialysis and electroporated (see section 2.2.5).

2.2.1.10  **Gibson assembly**
Gibson assembly reactions were prepared by combining 50 ng backbone and insert at a 6- or 9-fold molar ratio in a 20 µL reaction at 1X Gibson Assembly Master Mix (NEB, Hitchin, UK) and incubating for 1 h at 50 °C. 5 µL of the Gibson assembly reaction was transformed via chemical transformation (see section 2.2.4) or desalted via drop dialysis and electroporated (see section 2.2.5).

2.2.1.11  **Total DNA extraction**
Cell pellets were resuspended in 250 µL P1 buffer (QIAgem, Manchester, UK) with RNase at 100 µg / mL. SDS was added to 1 % and samples were vortexed briefly and incubated at room temperature for 5 min. Maxtract High-Density tubes (QIAgem, Manchester, UK) were centrifuged at 13,800 x g for 5 min and cell lysate was added
together with 275 µL of phenol : chloroform : isoamyl alcohol mix (25:24:1), mixing by inverting the tube 6 times. Tubes were then centrifuged at 13,800 x g for 5 min at 4 °C and the upper aqueous phase was transferred to 1.5 mL Eppendorf tubes containing 28 µL of sodium chloride (2 M). The tube was mixed by inverting 6 times. 758 µL ice-cold ethanol 100% was added and samples were mixed again by inversion, then placed at -80 °C for at least 30 min. Tubes were centrifuged at 13,800 x g for 15 min at 4 °C to pellet the DNA. The supernatant was decanted, and the white DNA pellet was washed with 1 mL 70 % ethanol followed by another centrifugation at 13,800 x g for 5 min at 4 °C. The supernatant was discarded again, and DNA pellets were dried for 1 h at 37 °C. Tubes were centrifuged briefly and 100 µL Tris-EDTA buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0) was added. Pellets were dissolved at 50 °C up to 1 h and stored at either 4 °C or –20 °C. If the DNA pellets were difficult to dissolve, more Tris-EDTA buffer was added.

2.2.1.12 Field-Inversion Gel Electrophoresis (FIGE)

The cell pellet was resuspended in one-half the final volume of plugs with cell suspension buffer (10 mM Tris, pH 7.2, 20 mM NaCl, 50 mM EDTA) and combined with an equal volume of melted 2% CleanCut agarose (Bio-Rad, Watford, UK), then equilibrated at 50 °C. The mix was transferred to CHEF disposable plug moulds (Bio-Rad, Watford, UK) and placed at 4 °C until solidified. The plugs were then placed in lysozyme buffer (10 mM Tris, pH 7.2, 50 mM NaCl, 0.2% sodium deoxycholate, 0.5 % sodium lauryl sarcosine, 1 mg / mL lysozyme) and incubated at 37 °C for 1 h. The plugs were then washed with wash buffer (20 mM Tris, pH 8.0, 50 mM EDTA) with gentle agitation at room temperature for 30 min and then transferred to proteinase K reaction buffer (100 mM EDTA, pH 8.0, 0.2% sodium deoxycholate, 1% sodium lauryl sarcosine, 1 mg / mL proteinase K) and incubated at 50 °C overnight. The plugs were washed four times with wash buffer and stored at 4 °C until use. If the plugs were to be digested with a restriction enzyme, the third wash was performed with 1 mM PMSF and the fourth with restriction enzyme.
buffer for 1 h. All buffers were used in a proportion of 25 mL per mL plugs. For restriction digestion of genomic plugs, fresh restriction buffer was added together with the enzyme and incubated at the conditions recommended by the manufacturer. Plugs were then washed once in wash buffer and stored at 4 °C until use.

Agarose gels were prepared in 1X Tris-borate-EDTA buffer (1 M Tris, 0.69 M boric acid, 0.01 M EDTA) at 1% Pulse-Field Certified agarose (Bio-Rad). The plugs were placed in the teeth of the gel comb prior to pouring the agarose mixture to set. FIGE was performed in a cold room in 0.5% TBE buffer using a FIGE Mapper (Bio-Rad, Watford, UK) running for 16 h with a forward voltage of 180 V, a reverse voltage of 120 V and a linear switch time ramp of 0.1 – 2 s. Gels were then stained in 3X GelRed solution at 0.1 M sodium chloride for 30 min and visualised using an Amersham 600 Imager (GE Healthcare, Amersham, UK).

2.2.1.13 PicoGreen assay
The DNA sample was dissolved to an appropriate concentration in Tris-EDTA buffer (10 mM Tris-HCl, 1 mM disodium EDTA, pH 8.0) and 100 µL were dispensed in triplicate in a clear flat-bottom black 96 well plate. λ DNA at a known concentration was diluted to make a standard in each plate. 100 µL QuantIt PicoGreen reagent (Thermo Fisher Scientific, Basingstoke, UK) was dispensed to each well, followed by a 5 min incubation at room temperature for 5 min. Fluorescence was measured at Ex485 / Em512 with a TECAN Infinite M200 plate reader (TECAN, Reading, UK).

2.2.2 Preparation of chemically competent cells
10 mL LB with the appropriate antibiotic in a 50 mL tube were inoculated with a colony of *E. coli* and grown from 16 – 20 h in an orbital shaker at 37 °C and 250 rpm. The following morning, a baffled 1 L shake flask with 200 mL LB with the appropriate antibiotic was inoculated with 2 mL of the overnight culture and grown in the same conditions until exponential phase was reached (OD₆₀₀ = 0.6 – 0.8). The culture was split into four 50 mL
tubes and centrifuged for 10 min at 4 °C and 1,500 x g. The supernatant was discarded, the pellets kept on ice and washed in 10 mL ice-cold CaCl₂ (100 mM) by gently pipetting up and down. The pellets were washed twice more and resuspended in 1.25 mL to which 200 µL 50% glycerol was added. 100 µL aliquots were made and stored at – 80 °C.

2.2.3 Preparation of electrocompetent cells

This protocol is nearly identical to that of chemically competent cell preparation (see section 2.2.2), with the exception that the three washes were performed with an ice-cold 10% glycerol solution.

2.2.4 Chemical transformation

An aliquot of chemically competent cells was thawed on ice for 20 min. The ice-cold DNA to be transformed was added to the aliquot and incubated on ice for 30 min, after which it was heat-shocked for 30 s at 42 °C and then placed on ice for 5 min. 1 mL of SOC media was added and the cells were allowed to recover at 37 °C 250 rpm from 1 h to overnight, then plated on LB or M9 agar with the appropriate antibiotic and incubated overnight at 37 °C.

2.2.5 Electroporation

An aliquot of electrocompetent cells was thawed on ice for 20 min. The ice-cold DNA to be transformed was added to the aliquot and transferred to an ice-cold MicroPulser® electroporation cuvette (Bio-Rad, Watford, UK) kept on ice. Electroporation was performed in a MicroPulser Electroporator (Bio-Rad, Watford, UK) with the appropriate settings for E. coli and the electroporation cuvette gap. The reaction was immediately transferred to a tube with 1 mL SOC media and the cells were allowed to recover at 37 °C 250 rpm from 1 h to overnight, then plated on LB or M9 agar with the appropriate antibiotic and incubated overnight at 37 °C.
2.2.6 Plasmid design and assembly

All oligonucleotides mentioned in this section can be found in the Appendix, section 8.2. PCR conditions can be found in the Appendix, section 8.3. All plasmids assembled in this work are summarised in Table 2-3.

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Size (bp)</th>
<th>Antibiotic resistance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQR1866</td>
<td>5385</td>
<td>Amp</td>
<td>pTTQ18 with the exonuclease III gene from <em>E. coli</em> cloned in the <em>Sma</em>I site under the <em>tac</em> promoter</td>
</tr>
<tr>
<td>pQR1867</td>
<td>5259</td>
<td>Amp</td>
<td>pTTQ18 with the λ exonuclease gene from phage λ cloned in the <em>Sma</em>I site under the <em>tac</em> promoter</td>
</tr>
<tr>
<td>pQR1868</td>
<td>5454</td>
<td>Amp</td>
<td>pTTQ18 with the T5 exonuclease gene from phage T5 cloned in the <em>Sma</em>I site under the <em>tac</em> promoter</td>
</tr>
<tr>
<td>pQR1869</td>
<td>6078</td>
<td>Amp</td>
<td>pTTQ18 with the exonuclease III and λ exonuclease genes cloned in the <em>Sma</em>I site under the <em>tac</em> promoter</td>
</tr>
<tr>
<td>pQR1870</td>
<td>6273</td>
<td>Amp</td>
<td>pTTQ18 with the exonuclease III and T5 exonuclease genes from <em>E. coli</em> cloned in the <em>Sma</em>I site under the <em>tac</em> promoter</td>
</tr>
<tr>
<td>pQR1871</td>
<td>6486</td>
<td>Cam</td>
<td>Controller plasmid from the dual recombinase expression system with the exonuclease expression cassette from pQR1870 in the Xhol-<em>Sma</em>I</td>
</tr>
<tr>
<td>pQR1872</td>
<td>6396</td>
<td>Amp</td>
<td>Expression plasmid from the dual recombinase expression system with the <em>Pvu</em>II restriction enzyme gene between the TP901 att sites replacing mkate2</td>
</tr>
<tr>
<td>pQR1873</td>
<td>6999</td>
<td>Amp</td>
<td>Expression plasmid from the dual recombinase expression system with the HpaII restriction enzyme gene between the TP901 att sites replacing mkate2</td>
</tr>
<tr>
<td>pQR1874</td>
<td>6228</td>
<td>Amp</td>
<td>Expression plasmid from the dual recombinase expression system with the ccdB gene between the TP901 att sites replacing mkate2</td>
</tr>
<tr>
<td>pQR1875</td>
<td>6357</td>
<td>Amp</td>
<td>Expression plasmid from the dual recombinase expression system with the denA gene between the TP901 att sites replacing mkate2</td>
</tr>
<tr>
<td>pQR1876</td>
<td>4816</td>
<td>Tet</td>
<td>pACYC184 with a constitutive expression cassette with sfGFP under promoter J23119 cloned in the <em>EcoRI-Nco</em>I site</td>
</tr>
</tbody>
</table>

(continued)
Table 2-3 (continued)

<table>
<thead>
<tr>
<th>Plasmid name</th>
<th>Size (bp)</th>
<th>Antibiotic resistance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQR1877</td>
<td>10693</td>
<td>Cam</td>
<td>pQR1871 with the sfGFP expression cassette from pQR1876 cloned in the Smal site</td>
</tr>
<tr>
<td>pQR1878</td>
<td>15198</td>
<td>Cam</td>
<td>pQR1877 with part of the Pseudomonas meta pathway cloned in the Smal-BsrGI site under Anderson collection promoter J23119</td>
</tr>
<tr>
<td>pQR1879</td>
<td>10677</td>
<td>Cam</td>
<td>pQR1877 with chromoprotein tinsel purple cloned in the Smal-BsrGI site under Anderson collection promoter J23119</td>
</tr>
<tr>
<td>pACYC184-NotIM</td>
<td>5030</td>
<td>Tet</td>
<td>pACYC184 with the NotI methyltransferase cloned in the EcoRI-NcoI site under the chloramphenicol promoter</td>
</tr>
<tr>
<td>pACYC184-PvuII</td>
<td>4979</td>
<td>Tet</td>
<td>pACYC184 with the PvuII methyltransferase cloned in the EcoRI-NcoI site under the chloramphenicol promoter</td>
</tr>
<tr>
<td>pACYC184-BamHIM</td>
<td>5240</td>
<td>Tet</td>
<td>pACYC184 with the BamHI methyltransferase cloned in the EcoRI-NcoI site under the chloramphenicol promoter</td>
</tr>
<tr>
<td>pACYC184-HpallIM</td>
<td>5045</td>
<td>Tet</td>
<td>pACYC184 with the Hpall methyltransferase cloned in the EcoRI-NcoI site under the chloramphenicol promoter</td>
</tr>
<tr>
<td>pPVU1t</td>
<td>5961</td>
<td>Tet</td>
<td>P. vulgaris restriction modification plasmid pPVU1 with a tetracycline resistance cassette in the EcoRI site</td>
</tr>
<tr>
<td>pPVU1t2</td>
<td>5486</td>
<td>Tet</td>
<td>pPVU1t plasmid with the restriction enzyme PvuII gene deleted and a PvuII site introduced instead</td>
</tr>
</tbody>
</table>

2.2.6.1 pQR1866, pQR1867, pQR1868

pTTQ18 was digested with Smal, dephosphorylated and purified with the DNA Clean and Concentrator kit (Zymo Research). Exonuclease III, λ exonuclease and T5 exonuclease were amplified from genomic DNA preparations of E. coli DH5α, and λ and T5 bacteriophage respectively. Oligonucleotides were designed to add a 21 - 25 bp overlapping sequence with pTTQ18, and a stop codon to interrupt the lacZα cistron, an RBS (AGGAGA) and 6 nucleotides upstream of the start codon. The oligonucleotides used were: #1-01, #1-02 for exonuclease III; #1-03, #1-04 for λ exonuclease; and #1-05, #1-06 for T5 exonuclease (sequence available in the Appendix, section 8.2.2). The
amplicons were either purified with a PCR purification kit or gel-extracted. The plasmids were assembled via CPEC, transformed and plated on LB Amp agar and incubated overnight at 37 °C. Recombinant colonies were screened by colony PCR using oligonucleotides #1-07 and #1-08, which were also used for sequencing (sequence available in the Appendix, section 8.2.2). The resulting plasmids were named pQR1866 (exonuclease III), pQR1867 (λ exonuclease) and pQR1868 (T5 exonuclease).

2.2.6.2 pQR1869, pQR1870
pQR1869 and pQR1870 were assembled in a two-insert CPEC and encode exonuclease III and either λ exonuclease or T5 exonuclease respectively. pTTQ18 was prepared in the same manner as described in section 2.2.6.1. Oligonucleotides were designed to include a ribosome binding site (RBS) in between coding sequences (CDS); 28 - 25 bp homology with either side of the SmaI site in pTTQ18; and 22 - 27 bp homology between inserts. The exonuclease III insert for pQR1869 was PCR-amplified with oligonucleotides #1-01 and #1-09; the exonuclease III insert for pQR1870 was PCR amplified with oligonucleotides #1-01 and #1-10; λ exonuclease, with #1-11 and #1-04; and T5 exonuclease, with #1-12 and #1-06. The plasmids were assembled via CPEC, transformed and plated on LB Amp agar and incubated overnight at 37 °C. Recombinants screened by colony PCR using oligonucleotides #1-07 and #1-08, which were also used for sequencing (sequence available in the Appendix, section 8.2.2).

2.2.6.3 Cloning restriction enzymes in pTTQ18
The restriction enzyme CDSs were PCR-amplified with oligonucleotides #2-05 and #2-06 (NotI); #2-07 and #2-08 (BamHI); #2-09 and #2-10 (PvuII); and #2-11 and #2-12 (HpaII). The amplicons were gel-purified and combined in a CPEC reaction with SmaI-digested and phosphorylated pTTQ18. 5 µL of the CPEC reaction were transformed into E. coli TOP10 cells, which were then plated on LB Amp agar and incubated overnight at 37 °C. The recombinants were screened by colony PCR using oligonucleotides #1-07
and #1-08, which were also used for sequencing (sequence available in the Appendix, section Table 8-28.2.2).

### 2.2.6.4 Cloning restriction enzymes in pQR445

The *Pvu*II CDS was PCR-amplified with oligonucleotides #2-01 and #2-02 (sequence available in the Appendix, section 8.2.3) to add an RBS, six spacer nucleotides and 20–35 bp homology with pQR445. pQR445 was digested with *BamHI* and dephosphorylated. Insert and vector were combined in CPEC reaction and 5 µL were transformed into TOP10 cells. The transformation was plated on LB Km agar and incubated overnight at 37 °C.

### 2.2.6.5 Cloning methyltransferases in pACYC184

pACYC184 was digested with *EcoRI* and *NcoI* for 2 h and gel-purified. The *NotI*, *Pvu*II, *BamHI* and *HpaII* methyltransferase genes were PCR-amplified using oligonucleotides #2-15 and #2-16; #2-17 and #2-18; #2-19 and #2-20; and #2-21 and #2-22 respectively from plasmids containing the CDSs (sequence available in the Appendix, section 8.2.3). Plasmids were assembled via CPEC, transformed in TOP10, and cells were plated on LB Tet agar and incubated overnight at 37 °C. Colonies were picked and plasmid DNA isolated and sent to sequencing using oligonucleotides #S-01 and #S-02 (sequence available in the Appendix, section 8.2.1).

### 2.2.6.6 pPVU1t and pPVU1t2

The DNA sequence encoding the tetracycline efflux protein expression cassette was PCR-amplified from pACYC184 using oligonucleotides #2-03 and #2-04 (sequence available in the Appendix, section 8.2.3). pPVU1 was obtained via plasmid DNA isolation of a *P. vulgaris* culture. Both plasmid and PCR amplicon were digested with *EcoRI* and ligated overnight. Positive colonies were screened by growth on tetracycline, and the resulting plasmid was termed pPVU1t. The *Pvu*II restriction enzyme CDS was then deleted by PCR-amplifying the vector at either side of the sequence, at the same time introducing a *Pvu*II site, using oligonucleotides #2-13 and #2-14 (sequence available in
the Appendix, section 8.2.3). The amplicon was purified, digested with *Pvu*II, ligated overnight at 16 °C, purified and electroporated. Cells were plated on LB Km and incubated overnight at 37 °C. The resulting plasmid was termed pPVU1t2.

**2.2.6.7 pQR1871**

Oligonucleotides #2-23 and #2-24 (sequence available in the Appendix, section 8.2.3) were designed in order to PCR-amplify the exonuclease and lacI expression cassettes from pQR1870 introducing *Xho*I and *Sma*I sites at either side and 10 bp to allow for restriction enzyme binding. The PCR was purified, and in parallel to the dual recombinase system controller plasmid, digested with *Xho*I and *Sma*I and purified again. Insert and backbone were combined in a 2:1 ratio due to their similar size, ligated and transformed into competent *E. coli* DH5αZ1. Cells were plated on LB Cam agar and incubated overnight at 37 °C. Positive colonies were screened via miniprep digestion with *Bam*HI and *Xho*I.

**2.2.6.8 pQR1872, pQR1873, pQR1874, pQR1875**

Given the lack of useful restriction sites in the expression plasmid, the backbone was generated through PCR amplification with oligonucleotides #2-25 and #2-26, designed to introduce *Bam*HI and *Xho*I sites and exclude mkate2. A gradient PCR was performed to choose the best conditions, with an annealing temperature of 52 – 58 °C and a DMSO concentration of 0 – 5 %.

The sequences for *Not*I, *Bam*HI, *Pvu*II and *Hpall* were PCR amplified from genomic DNA extractions of *P. vulgaris*, *H. influenzae*, *N. otitidis-caviarum* and *B. amyloliquefaciens*. The sequences for ccdB and denA were ordered as gBlocks (sequence available in the Appendix, sections 8.4.1 and 8.4.2). For all genes, oligonucleotides were designed to introduce *Bam*HI and *Xho*I sites at either side of the CDS with 10 bp to facilitate restriction enzyme binding and an RBS at the 5’ end of the CDS via PCR: oligonucleotides #2-27, #2-28 for *Pvu*II; #2-29 and #2-30 for *Not*I; #2-31 and #2-32 for *Bam*HI; and #2-33 and #2-34 for *Hpall* (sequence available in the Appendix, section 8.2.3).
Each PCR-amplified insert, both gBlocks and the expression plasmid backbone were digested with XhoI and BamHI and purified with a PCR purification kit. Backbone and each insert were ligated and transformed into DH5αZ1. Cells were plated on LB Amp agar and incubated overnight at 37 °C. Positive colonies were screened via colony PCR using oligonucleotides #2-35 and #2-36 (sequence available in the Appendix, section 8.2.3). Confirmed assembled plasmids were termed pQR1872 (PvuII), pQR1873 (HpaII), pQR1874 (ccdB), and pQR1875 (denA).

2.2.6.9 pQR1876
pACYC184 was digested with EcoRI and NcoI and the backbone was purified via gel purification. A gBlock (sequence available in the Appendix, section 8.4.3) was designed containing promoter J23119 from the Anderson family of constitutive promoters, superfolder GFP (sfGFP) flanked by XhoI and BsrGI sites, and the rrnB T1 terminator, all flanked by the restriction sites and a 30 bp overlap with the backbone. Plasmid pQR1876 was then assembled via Gibson assembly. 5 µL of the reaction were electroporated into TOP10 cells, plated on LB Cam and incubated overnight at 37 °C. Positive colonies were screened via presence of green fluorescence.

2.2.6.10 pQR1877, pQR1878, pQR1879, pQR1880
Plasmid pQR1871 was Smal-digested, dephosphorylated and purified to prepare the backbone for cloning of the sfGFP expression cassette PCR-amplified from pQR1876 with oligonucleotides #3-01 and #3-02 (sequence available in the Appendix, section 8.2.4). Plasmid pQR1877 was then assembled via Gibson assembly (see section 2.2.1.10), chemically transformed into TOP10 cells, plated on LB Cam and incubated overnight at 37 °C. Positive colonies were selected via green fluorescence.

Plasmids pQR1878 (fragment of the Pseudomonas meta pathway), pQR1879 (tinsel purple) and pQR1880 (CV2025 transaminase) were originated from pQR1877. Given that the Pseudomonas meta pathway insert had three XhoI sites, the backbone was obtained by PCR-amplifying pQR1877 with oligonucleotides #3-03 and #3-04 to
introduce Smal and BsrGI sites instead. The pathway was PCR-amplified from plasmid pQR226 using oligonucleotides #3-05 and #3-06; tinsel purple was amplified from the tinsel purple plasmid from the Protein Paint Box collection (ATUM, California, UK) using oligonucleotides #3-07 and #3-08; and CV2025 transaminase was amplified from pQR801 (Kaulmann et al., 2007) using oligonucleotides #3-09 and #3-10. Oligonucleotide sequences are available in the Appendix, section 8.2.4. All inserts and backbone were digested with Smal at 25 °C for 1 h, then with BsrGI 37 °C for 2 h. The digestions were purified, and assemblies were performed in a cycled ligation protocol (see 2.2.1.8). Ligations were chemically transformed into TOP10 cells, plated on LB Cam and incubated overnight at 37 °C. Colonies were screened as explained in section 2.2.7.1. Positive pQR1879 (tinsel purple) colonies were screened via presence of a purple colour, and positive pQR1868 (meta pathway) and pQR1880 (CV2025 TAm) colonies were screened as explained in section 2.2.7.2.

2.2.6.11 pQR1886
A gBlock was designed containing the first 134 bp of the CV2025 TAm sequence, flanked by Smal and MauBI sites (sequence available in the Appendix, section 8.4.4). Both gBlock and pQR1880 were digested with Smal and MauBI, purified and ligated in a cycled ligation protocol. 5 µL of the ligation were electroporated into TOP10 cells and plated on LB Cam agar and incubated overnight at 37 °C. Positive colonies were screened by sequencing using oligonucleotides #S-11 and #S-12 (sequence found in the Appendix, section 8.2.4).

2.2.7 Colony screening

2.2.7.1 Catechol 2,3-dioxygenase screening
Plates with colonies potentially expressing part of the Pseudomonas meta pathway were sprayed with 100 mM catechol and allowed to dry at room temperature. Positive colonies were screened through the yellow colour they acquired through production of 2-hydroxymuconic aldehyde.
2.2.7.2 **CV2025 transaminase (TAm) screening**

The same method used by Baud *et al.* (2015) was used for screening of CV2025 TAm transformants. Briefly, replicates of plates with colonies potentially expressing the CV2025 TAm were made; one plate was sprayed with 5 mM benzaldehyde and 12.5 mM 2-(4-nitrophenyl)ethan-1-amine hydrochloride and incubated at room temperature for 1 h. Positive colonies turned red and were picked from the other plate.

2.3 **Protein expression studies**

2.3.1 **Protein extract preparation**

Pellets were thawed on ice, resuspended in 1 mL Tris-HCl buffer (pH = 7.5) and sonicated on ice for 4 cycles, 5 s on and 10 s off. Protein extracts were centrifuged for 5 min at 13,800 x g and 4 °C and the supernatant consisting of soluble protein was transferred to a new tube and kept on ice.

2.3.2 **Protein activity assays**

2.3.2.1 *Exonuclease activity assays*

8 µL of protein extract was added to 32 µL of *EcoRV, BpmI* or *HindIII*-digested λ DNA at 5 ng/mL, vortexted briefly and placed in a thermomixer at 37 °C. 10 µL were sampled and immediately mixed with 3 µL loading dye (NEB) at 1, 5, 15 and 30 min. 10 µL of digested λ DNA with no protein extract added was used as the 0 min timepoint. Samples were loaded in a 1% agarose TBE gel with 1X GelRed nucleic acid stain (Biotium, California, US) and ran at 95 V for 1 h. Gel images were taken with an Amersham 600 Imager (GE Healthcare, Amersham, UK).

2.3.2.2 *Endonuclease activity assays*

To test activity of *PvuII* and *HpaII*, 0.5 µL λ DNA (NEB, Hitchin, UK) were combined with 5 µL clarified lysate, 1.5 µL 3.1 or CutSmart buffer (NEB, Hitchin, UK) and topped up to 15 µL with molecular-grade water. To test activity of denA, 0.5 µL λ DNA (NEB, Hitchin, UK) were combined with 5 µL clarified lysate in 1x reaction buffer (10 mM Tris-HCl, pH
7.5; 5 mM MgCl₂, 10 mM NaCl, 0.1 mM Na₂EGTA (adapted from Carlson, Kosturko and Nystrom (1999)). Assays were incubated at 37 °C for 1 h, and then 3 µL purple loading dye (NEB, Hitchin, UK) was added and the whole reaction was run on a 1% agarose TBE gel along with the appropriate controls.

2.3.2.3 Transaminase assay
The transaminase activity assay is that of Baud et al. (2015). Protein extracts were prepared as described in section 2.3.1, but in phosphate buffer (100 mM, pH = 7.5) in this case.Briefly, transaminase assays were performed in a total volume of 200 µL containing PLP (0.2 mM), amine acceptor (benzaldehyde at 10 mM) and amine donor (2-(4-nitrophenyl)ethan-1-amine hydrochloride at 25 mM) in phosphate buffer with 50 µL protein extract. Each extract was assayed in triplicate in a 96 well plate alongside the appropriate controls (no protein extract; no amine acceptor; and no amine donor). Plates were sealed and incubated at 30 °C and 200 rpm in an orbital shaking incubator for 48 h.

2.3.2.4 Meta pathway and catechol 2,3-dioxygenase assays
A volume of cells equivalent to 1 mL at an OD₆₀₀ = 0.5 was harvested by centrifugation at 3,500 x g for 10 min, the supernatant was discarded, and the pellet was resuspended in 1 mL PBS, which was then split in two 1.5 mL tubes. In order to assay catechol 2,3-dioxygenase activity, catechol was added to 5 mM to one of the tubes, the cells were vortexed briefly and incubated for 10 min at room temperature. As control, PBS was added instead to the other tube. Cells were incubated at 90 °C for 5 min to stop the reaction. In order to assay the whole fragment of the meta pathway, sodium benzoate was added to 1 mM instead and incubated at room temperature overnight. The amount of 2-hydroxymuconic aldehyde formed was measured by absorbance at 395 nm with a spectrophotometer. Prior to each measurement, the spectrophotometer was blanked with the control cell sample with PBS.
2.3.3 SDS-PAGE

For Tris-glycine gels, cell pellets were resuspended in 100 µL Laemmli buffer with 50 mM DTT and incubated at 95 °C for 5 min. 10 µL sample were loaded in 4 - 20% mini-PROTEAN TGX precast gels (Bio-Rad) and ran in Tris-Glycine buffer for 40 min at 200 V. Gels were then stained for at least 30 min in InstantBlue (Expedeon Ltd, Cambridge, UK).

For Bis-Tris gels, cell pellets were resuspended in 60 µL 1X LDS sample buffer with 100 mM DTT and incubated at 70 °C for 10 min. 10 µL sample were loaded in a NuPage 10% Bis-Tris gel (Thermo Fisher Scientific, Basingstoke, UK) and run for 1 h at 170 V in MOPS buffer. The protocol for Western blot was followed (see section 2.3.4).

2.3.4 Western blot for His-tagged protein detection

The excess SDS-PAGE gel was trimmed and transferred to a PVDF membrane for 7 min using a Trans-Blot Turbo Mini PVDF transfer pack and a Trans-Blot Turbo transfer system (Bio-Rad, Watford, UK). The membrane was blocked overnight by incubating with blocking buffer (5% skimmed milk in TBST buffer) with mild shaking at 4 °C. The membrane was then washed with TBST buffer 3 times for 5 min each, and then incubated for 2 h with mild shaking at room temperature with 10 mL antibody incubation buffer (2.5% skimmed milk in TBST buffer) and 1:1,250 rabbit anti-6X His-tag primary antibody (Abcam, Cambridge, UK). The membrane was washed again with TBST buffer 3 times for 5 min each, and then incubated for 1 h with mild shaking at room temperature with 10 mL antibody incubation buffer and 1:10,000 goat anti-rabbit IgG highly cross-adsorbed secondary antibody, horseradish peroxidase (Thermo Fisher Scientific, Basingstoke, UK). The membrane was washed 3 times and then 6 mL Pierce ECL Plus Western Blotting solution (Bio-Rad, Watford, UK) was added followed by a 10 min incubation with mild shaking at room temperature. The blot was then visualised with the
chemiluminescence function of an Amersham Imager 600 (GE Healthcare, Amersham, UK).

2.4 Experiments involving cell growth and analysis

2.4.1 Working cell banks for exonuclease expression growth curves
A colony of *E. coli* CSR603 harbouring plasmids pTTQ18, pQR1866, pQR1867, pQR1868, pQR1869 or pQR1870 was inoculated in 250 mL baffled flasks with 50 mL LB Amp and grown until OD$_{600}$ = 1.5. Cells were then resuspended in LB Amp with 25% glycerol to an OD$_{600}$ = 2, and 1 mL aliquots were prepared and stored at – 80 °C.

2.4.2 Growth curves for exonuclease overexpression
Working cell bank vials were thawed on ice for 20 min and 0.5 mL were used to inoculate 50 mL LB Amp in 250 mL baffled shake flasks, then grown for 14 – 16 h at 30 °C and 250 rpm in an orbital shaker. When the culture approached mid exponential phase, it was diluted in LB Amp to an initial OD$_{600}$ = 0.1, and 100 mL were grown in duplicate in 500 mL baffled shake flasks in an orbital shaker at 37 °C and 250 rpm for a total of 24 h. At 3 h cultures were induced at 0.5 mM IPTG where pertinent. OD$_{600}$ was recorded every hour, and at times 3, 4, 5, 7, 9, 12 and 24 h samples consisting of 1 mL and a culture equivalent for each time point (volume equivalent to 1 mL at an OD$_{600}$ = 0.5) were taken for DNA extraction and exonuclease activity assays respectively. Samples were centrifuged for 10 min at 1,500 x g, supernatant discarded, and pellets stored at – 20 °C until further analysis.

2.4.3 Determination of growth-inhibiting MMC and CI concentrations
A colony of *E. coli* DH5αZ1 or CSR603 harbouring plasmid pQR1870 was inoculated into 10 mL LB Amp in a 50 mL Falcon tube and grown at 37 °C and 250 rpm in an orbital shaker overnight. This was used to inoculate 200 mL LB Amp in duplicate in 1 L baffled shake flasks at an initial OD$_{600}$ = 0.05 and grown at the same conditions until OD$_{600}$ = 1.
10 mL of this culture were dispensed in duplicate onto each Falcon tube containing different quantities of either MMC, CI, or both and returned to the incubator. 0.5 mL samples were taken at 2, 4, 6 and 21 h for OD$_{600}$ measurements. An additional 0.5 mL sample was taken at 2 h to plate on LB agar and incubated at 37 °C for 24 h for cell viability determination. MMC and CI concentrations for each sample are specified in Chapter 4, Table 4-1.

2.4.4 Combined antibiotic addition and nuclease induction experiments

A colony of *E. coli* DH5αZ1 or CSR603 harbouring either plasmid pQR1870 or pQR1871 and pQR1872 was inoculated into 10 mL of LB Amp or LB Amp Cam and incubated in an orbital shaker for 16 - 24 h at 37 °C and 250 rpm. This starter culture was diluted to an initial OD$_{600}$ = 0.05 in 200 mL LB Amp Cam in two 1 L baffled shake flasks and grown at the same conditions until OD$_{600}$ = 0.5 - 0.8 (approximately 3 h for DH5αZ1 and 5 h for CSR603). One flask was then induced at 0.5 mM IPTG in the case of plasmid pQR1870, or 1% L-arabinose and 0.5 mM IPTG in the case of plasmids pQR1871 and pQR1872. The equivalent volume of PBS was added to the non-induced flask. Incubation continued for 4 h. 50 mL were then added to 250 mL baffled shake flasks with different MMC and CI amounts in duplicate and returned to the incubator. Samples were taken after 3, 6 and 17 h. 1 mL samples were centrifuged at 3,500 x g for 10 min, supernatant discarded, and cell pellets stored at – 20 °C for DNA extraction. 0.5 mL samples were taken for OD$_{600}$ measurements. At 6 h, an additional 0.5 mL were taken for plating on LB agar and incubated at 37 °C for 24 h for cell viability determination.

2.4.5 DSB-inducing protein induction experiments

*E. coli* DH5αZ1 or CSR603 strains harbouring plasmid pQR1871 or the controller plasmid, and either pQR1872, pQR1873, pQR1874, pQR1875 or the empty expression plasmid were obtained by sequential transformation. A colony was inoculated into 10 mL LB Amp Cam and incubated in an orbital shaker for 16 - 24 h at 37 °C and 250 rpm. This
starter culture was diluted to an initial \(OD_{600} = 0.05\) in 50 mL Amp Cam in three 250 mL baffled shake flasks and grown at the same conditions until the time of induction, when L-arabinose and IPTG were added at 1 % and 0.5 mM IPTG if necessary. If not induced, the equivalent volume of PBS was added to the non-induced flask. Samples consisted of: 0.5 mL for \(OD_{600}\) measurements; 1 mL for total DNA extraction, for which samples were centrifuged at 3,500 x g for 10 min, supernatant discarded, and cell pellets stored at –20 °C until analysis; and 250 µL for viability count (see section 2.4.6).

2.4.6 Serial dilutions and colony-forming unit (cfu) determination

Serial dilutions were performed using a TECAN Freedom EVO (TECAN, Reading, UK). 250 µL of cell culture was first manually transferred to a 96 deep-well plate and the pipetting robot performed serial dilutions with LB, plating 60 µL of each dilution in 6 well microplates with LB agar and antibiotic if pertinent and 3 glass beads. After dispensing, plates were manually shaken for 10 s and the glass beads were tipped into a beaker with Virkon solution (DuPont, Stevenage, UK). Plates were incubated for at least 16 h at 37 °C and colonies were counted from wells in which there were between 25 and 250 colonies.

2.4.7 Monitoring cell growth and fluorescence in a 96 well plate

Controller recombinase plasmid and the expression plasmid were transformed sequentially in *E. coli* DH5αZ1 chemically competent cells and plated on LB agar Amp Cam. A colony was picked and grown overnight in 10 mL LB Amp Cam. The following day, 100 µL of this overnight were used to inoculate 10 mL LB Amp Cam and grown for 3 h. 200 µL of this culture were dispensed on a black 96 well plate with a flat transparent bottom (Greiner) with different quantities of L-arabinose, to 0, 0.01, 0.1 and 1 %. Each condition was tested in triplicate. The plate was covered with a Breathe-Easy sealing membrane (Sigma) and transferred to a TECAN Infinite 200 PRO plate reader (TECAN, Reading, UK) at 37 °C. The plate was shaken every 15 min for 10 s during a total of 13 h,
after which the measurements of absorbance at 600 nm and fluorescence at Ex588 / Em633 were performed.

2.4.8 Genome ablation and UV irradiation experiments

*E. coli* CSR603 pQR1871 pQR1872 were centrifuged for 15 min at 1,500 x g and 4 °C and resuspended in PBS at an OD_{600} = 0.2. 12.5 mL were transferred to a Petri dish with an ethanol-sterilised magnetic stirrer and irradiated at different times with a CL-1000 Ultraviolet Crosslinker (Fisher Scientific, Loughborough, UK) placed on a magnetic stirrer to enable constant mixing of the liquid suspension. This crosslinker emits light at a wavelength of 254 nm. The preset factory settings are 2 min and 120,000 µJ / m²; therefore, it can be estimated that UV irradiation using this crosslinker occurs at a fluence rate of 10 W / m², or 10 J / m²s. 10 mL of the irradiated suspension were centrifuged for 15 min at 1,500 x g and 4 °C, resuspended in 10 mL LB and incubated at 37 °C and 180 rpm in an orbital shaking incubator. After 1 h, ceftriaxone was added to 100 µg / mL and incubation proceeded for another 16 h. 3 mL were plated on LB agar to ensure there were no survivors.

When performing this experiment to evaluate activity of constitutively expressed proteins, 5 mL of the sample were centrifuged and resuspended in 650 µL PBS. For sfGFP fluorescence quantification, each pellet was resuspended in 650 µL PBS and 200 µL were transferred in triplicate to black 96 well plates with clear bottom. To test catechol 2,3-dioxygenase activity, the assay was performed as explained in section 2.3.2.4 and 200 µL were transferred in triplicate to clear flat bottom 96 well plates. Spectrophotometric readings were performed a TECAN Infinite 200 PRO plate reader (TECAN, Reading, UK): absorbance at 600 nm and fluorescence at Ex485 / Em512 for sfGFP quantification, and absorbance at 600 and 395 nm for 2-hydroxymuconic aldehyde quantification. CV2025 TAm activity tests were performed as described in 2.3.2.3.
2.4.9 Cell staining

Genome-ablated (GA) cells were stained along different controls: non-induced, both non-UV irradiated and UV-irradiated; induced, not UV-irradiated; and live and dead CSR603 pQR1871 pQR1872 cells. Dead cells were obtained by resuspension in 70 % isopropanol. Since two stains are used, unstained and Fluorescence-Minus-One (FMO) controls were performed for all samples. All centrifugation steps were performed at 860 x g for 10 min, and all washes are performed in 1 mL PBS.

A cell suspension volume equivalent to $3.2 \cdot 10^8$ cells was harvested via centrifugation at 860 x g for 10 min, washed once and resuspended in PBS. 1 µL of Live-or-dye (488/515) stain (Biotium, California, US) was added, the cell suspension was immediately mixed and incubated for 30 min at room temperature protected from light. Cells were then washed twice and fixed via resuspension in 1 mL PBS with 2 % paraformaldehyde for 30 min at room temperature. Two washes were performed, and cells were resuspended in 1 mL PBS to which Hoechst 34580 (Thermo Fisher Scientific, Basingstoke, UK) was added at a final concentration of 6 µg / mL. Cells were incubated at room temperature for 30 min protected from light, washed once and resuspended in PBS. Samples were then analysed via flow cytometry and fluorescence microscopy.

2.4.10 Flow cytometry

Flow cytometry was performed in a BD FACSJazz Cell Sorter (BD, Wokingham, UK). Lasers were calibrated using SPHERO Rainbow Calibration particles (BD, Wokingham, UK). For transformant screening, 10 million events were collected; for stained samples as described in section 2.4.9, a minimum of 50,000 events were collected. Results were analysed with FlowJo V.10 (BD).

2.4.11 Fluorescence microscopy

Samples were mounted in 1 % agarose microscope slides and observed in an inverted epifluorescence Olympus IX81 microscope (Olympus Corporation, Tokyo, Japan)
equipped with an IX2-UCB control board and a 100x oil immersion objective. Hoechst 34580 and Live-or-Dye were excited at 405 nm and 488 nm respectively, and emission detected using the DAPI and FITC filter sets. Images were processed using the Fiji image processing package on ImageJ (Schindelin et al., 2012).
3 Chapter 3. Effect of exonuclease overexpression in *E. coli* CSR603

3.1 Introduction

In order to construct a genome ablation cassette that will completely digest the *E. coli* genome and plasmid DNA down to nucleotides *in vivo*, both endonucleolytic and exonucleolytic elements are needed. Whereas different agents can produce DSBs, only exonucleases are capable of DNA degradation. Exonucleases are enzymes that degrade the polynucleotide chain by hydrolysing phosphodiester bonds between nucleotides from the ends of a DNA fragment only. Their main role in bacteria is repair: for instance, *E. coli* exonuclease II and VI are the 3’ → 5’ and 5’ → 3’ proofreading activities of DNA polymerase I respectively. Bacteriophages also use exonucleases in their infection cycle in order to degrade host DNA down to nucleotides, using them to replicate the viral genome and assemble new virions. Some bacteriophages use host exonucleases, while others have them encoded in their genome and are transcribed and translated by the host machinery.

Exonucleases can be classified according to different variables, depending on the individual specificity of the enzyme: polarity of digestion (from the 3’ end to 5’ or from the 5’ end to 3’); ability to degrade ssDNA, dsDNA or both; activity without a phosphate group in 5’; ability to initiate digestion on nicked, blunt or DNA with overhangs; and extent of degradation (short oligonucleotides, ssDNA, dNMPs). [deleted] In this work, one 3’ → 5’ exonuclease was selected, exonuclease III of *E. coli*, a multifunctional enzyme with endonuclease, diesterase repair, exonuclease and RNAse H activities; and two 5’ → 3’ exonucleases, λ and T5 bacteriophage exonucleases, which degrade host DNA upon infection in order to scavenge nucleotides for viral replication. These exonucleases were selected based on ability to digest a wide range of DNA substrates down to dNMPs. These are expressed both individually and combined in *E. coli* CSR603, a Maxicell strain that has impaired DNA repair as mentioned in Chapter 1, section 1.4.5.2. The
exonuclease-encoding plasmids are illustrated in Figure 3-1. Growth curves, total DNA content and *in vitro* activity from protein extracts are compared in order to determine which exonuclease combination is most suitable for a genome ablation cassette.

Figure 3-1. Simplified plasmid map of the different exonuclease plasmids. A more detailed version of the plasmid maps can be found in sections 8.1.4 to 8.1.8 in the Appendix.
3.2 Results and discussion

3.2.1 Cloning single exonucleases in pTTQ18

The three exonucleases used in this work were selected based on availability of template DNA in the Ward lab. Exonuclease III, λ phage exonuclease and T5 phage exonuclease were PCR-amplified from their templates introducing overlapping regions with the destination vector pTTQ18 (plasmid map found in the Appendix, section 8.1.1), yielding the expected fragment size as seen in Figure 3-2, Figure 3-3 and Figure 3-4.

![Figure 3-2](image)

*Figure 3-2. Gel depicting PCR amplification of exonuclease III from the E. coli MG1655 genome, 872 bp (lane 1, blue arrow). This fragment was gel-purified given the primer dimerisation (lane 1, green arrow). 10 µL of the PCR reaction were loaded. The ladder (lane M) used is the 100 bp DNA ladder (NEB).*

![Figure 3-3](image)

*Figure 3-3. Gel depicting PCR amplification of λ exonuclease from the λ bacteriophage genome, 749 bp (lane 1, blue arrow). 10 µL of the PCR reaction were loaded. The ladder (lane M) used is the 100 bp DNA ladder (NEB).*
As explained in Chapter 2, section 2.2.6.1, the forward oligonucleotides used to amplify the exonuclease inserts (oligonucleotides #1-01, #1-03 and #1-05; sequence available in the Appendix, section 8.2.2) inserted 21 – 23 bp of homology with pTTQ18 in order to enable cloning via CPEC, a stop codon, an RBS, six spacer nucleotides and the required number of nucleotides homologous with the exonuclease genes. The pTTQ18 multicloning site is found within the lacZ subunit α CDS, so a stop codon in frame with lacZ subunit α was introduced to interrupt any unwanted translation immediately upstream of the exonuclease, resulting in a two-cistron system: 7 amino acids of lacZ subunit α, and the exonuclease. This is illustrated in Figure 3-5.
Figure 3-5. Illustration of the two-cistron system created in order to clone the exonucleases into vector pTTQ18. The procedure is illustrated with exonuclease III. pTTQ18 is first cleaved on the Smal site within the lacZα CDS, and then combined in a CPEC reaction with the PCR-amplified fragment containing the stop codon TAG (in red), the RBS, six spacer nucleotides and the exonuclease CDS. The assembled product contains two cistrons: one composed by the first 7 codons of lacZα, and the exonuclease gene.

After obtaining the desired inserts, CPEC was performed. As explained in Chapter 1, section 1.2.3, CPEC is a method of DNA assembly characterised for its simplicity, given that it only requires a DNA polymerase and vector assembly can be performed in a single PCR step with multiple fragments at the same time as long as there is a minimum overlap between fragments (Quan and Tian, 2009). In order to assemble a single gene into a vector, the original publication by Quan and Tian (2009) states a single PCR cycle is enough for successful cloning, and includes a comparison of an assembly after 1, 2 and 5 PCR cycles run in an agarose gel. They observed considerably more product after 5 cycles, and therefore the protocol used in this study includes 5 cycles. In addition, their protocol consisted of a polymerase extension step of 15 seconds per kb of final product.
However, common practice in the Ward group is to perform a longer extension step of 30 seconds per kb, which is what was done in this study. A comparison of both variations of extension time was made and run in a gel as seen in Figure 3-6: more product is observed in lane 3, a CPEC reaction with 30 seconds per kb, compared to lane 2, with 15 seconds per kb, since the bands appearing above the linearised plasmid band become brighter. Another indication of the assembly occurring is the brightness of the insert band, which becomes lighter as it is being integrated into the linearised backbone.

![Figure 3-6. Gel depicting the CPEC reaction for assembly of the exonuclease III insert into the pTTQ18 backbone. 10 μL of each reaction were loaded. Lane 1 is a control consisting of both inserts (pTTQ18 backbone, blue arrow, 4563 bp; and insert, green arrow, 872 bp) combined, with no assembly; lane 2 is a CPEC reaction with 15 seconds per kb final product; lane 3 is a CPEC reaction with 30 seconds per kb final product. In lanes 2 and 3, newly assembled plasmid is highlighted with the orange arrow. The ladder (M) used is the 1 kb Extend DNA ladder (NEB).]

After CPEC, assemblies were transformed into *E. coli* TOP10. Several dilutions of the transformation were plated, and colonies were picked from the lowest dilution as it yielded a few distinct colonies. Positive colonies were screened via colony PCR as depicted in Figure 3-7 and Figure 3-8.
Figure 3-7. Gel depicting colony PCR after CPEC assembling pQR1866. The whole colony PCR reactions (10 µL) were loaded. Lane 1 is a control consisting of a PCR with empty pTTQ18 as template (500 bp band, green arrow; the band at 3 kb is the plasmid template, orange arrow). Positive colonies yield a 1322 bp band (lanes 4, 6, 9, 10, 11; blue arrows); negative colonies yield a 500 bp band (lanes 2, 3, 5, 7; green arrows). No clear band is observed in lane 8, probably due to an incomplete resuspension of the colony when picking it. The ladder (M) used is GeneRuler DNA Ladder Mix (NEB).

Figure 3-8. Gel depicting colony PCR after CPEC assembling pQR1868 (lanes 1-9) and pQR1867 (lanes 10-12). The whole colony PCR reactions (10 µL) were loaded. Positive colonies yield a 1196 bp (blue arrow) and a 1391 bp band (green arrow) respectively (lanes 1, 6, 7, 8, 9, 10, 11, 12). In lane 3, the presence of multiple bands indicates that two different colonies were probably picked (orange square). The bands observed in lanes 2, 4 and 5 belong to negative colonies (500 bp, orange arrow). The ladder (M) used is GeneRuler DNA Ladder Mix (NEB).

In the lowest dilution plate, a total of 9 colonies were obtained for pQR1866, 3 for pQR1867, and 9 for pQR1868. After the screening, the percentage of positive assemblies was found to be 50 %, 100 % and 44.4 % for pQR1866, pQR1867 and pQR1868 respectively; the high empty plasmid background is likely due to an incomplete digestion or dephosphorylation of pTTQ18 and recircularisation, since it was linearised by digestion in a single site as part of the cloning procedure. The plasmid maps for pQR1866, pQR1867 and pQR1868 can be found in the Appendix, sections 8.1.4 to 8.1.6.
3.2.2 Cloning exonuclease pairs in pTTQ18

The next step was to assemble two plasmids with two exonuclease pairs, consisting of a 3’ → 5’ (exonuclease III) and a 5’ → 3’ exonuclease (λ or T5 exonucleases): pQR1869 (exonuclease III + λ exonuclease) and pQR1870 (exonuclease III + T5 exonuclease).

The plasmid maps for pQR1869 and pQR1870 can be found in the Appendix, sections 8.1.7 and 8.1.8. This was necessary to ensure DNA digestion was complete down to nucleotides, since different exonucleases work best from one of the two polarities. To obtain these plasmids, instead of using pQR1866, cloning was done from scratch using pTTQ18: as explained in Chapter 2, section 2.2.6.2, overlapping regions with pTTQ18 and with each other fragment were introduced via PCR (as explained in section 3.2.1 and in Chapter 2, section 2.2.6.2), and the PCR-amplified fragments are shown in Figure 3-9.

![Figure 3-9. Gel depicting PCR amplification of the inserts (blue arrows) needed for pQR1869 (exonuclease III and λ exonuclease, lanes 1-3 and 4-6, 868 and 724 bp respectively) and pQR1870 (exonuclease III and T5 exonuclease, lanes 7-9 and 10-12, 860 and 922 bp respectively). The larger bands (orange arrows) corresponds to the plasmid DNA template (pQR1866, pQR1867 and pQR1868), since too much was loaded in this PCR. 10 µL of each PCR reaction were loaded in each lane. Inserts were gel-purified. The ladder (M) used is GeneRuler DNA Ladder Mix (NEB). The low quality of the gel is due to the need to increase brightness and contrast in order to see all gel fragments.](image)

Colonies were screened in the same manner as for plasmids pQR1866, pQR1867 and pQR1868 and the resulting colony PCR gel can be seen in Figure 3-10. Given the previous high proportion of empty plasmid yielding many negatives, pTTQ18 was prepared again with more restriction enzyme and the cloning was shown to be successful in all 6 colonies picked for each plasmid.
3.2.3 Exonuclease expression

3.2.3.1 SDS PAGE

SDS PAGE was run in order to verify that the exonucleases were being expressed as expected (Figure 3-11, Figure 3-12 and Figure 3-13). To obtain the samples, cultures of *E. coli* CSR603 with pQR1866, pQR1867 or pQR1868 were grown to early exponential phase (OD$_{600} = 0.6$), induced with IPTG and incubated overnight. Whole-cell samples were run. Different IPTG concentrations (0.1, 0.5 and 1 mM) were run to ascertain whether inducer concentration made a difference, alongside a non-induced control.
Figure 3-11. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1866. Lane 1 is an uninduced control; lane 2 is an induced sample at 0.1 mM IPTG; lane 3, at 0.5 mM; lane 4, at 1 mM. The highlighted band corresponds to exonuclease III (31 kDa). The ladder used (M) is the PageRuler Unstained Protein Ladder (ThermoFisher).

Figure 3-12. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1867. Lane 1 is an uninduced control; lane 2 is an induced sample at 0.1 mM IPTG; lane 3, at 0.5 mM; lane 4, at 1 mM. The highlighted band corresponds to λ phage exonuclease (26 kDa). The ladder used (M) is the PageRuler Unstained Protein Ladder (ThermoFisher).

Figure 3-13. SDS PAGE of whole-cell samples of E. coli CSR603 pQR1868. Lane 1 is an uninduced control; lane 2 is an induced sample at 0.1 mM IPTG; lane 3, at 0.5 mM; lane 4, at 1 mM. The highlighted area is where T5 phage exonuclease (33.5 kDa) was expected to appear. The ladder used (M) is the PageRuler Unstained Protein Ladder (ThermoFisher).
The bands corresponding to both exonuclease III and λ exonuclease can be clearly seen at all IPTG concentrations in Figure 3-11 (31 kDa) and Figure 3-12 (26 kDa) respectively, and there does not seem to be an obvious correlation between IPTG concentration and amount of protein expressed overall. However, no obvious band can be observed for T5 exonuclease in Figure 3-13 (33.5 kDa). This does not mean that the protein is not being expressed: it is possible that this protein is being expressed at such low levels that a simple protein gel does not display it, and perhaps alternative methods like a Western blot would be more conclusive. However, since presence of T5 exonuclease can be tested via activity in vitro (which is done in the next section), this was not explored any further.

### 3.2.3.2 Exonuclease activity in vitro

To test activity in vitro, a simple assay consisting of adding protein extract of an induced culture to λ DNA digested with EcoRV. Methods of protein extract preparation and activity assay are explained in Chapter 2, sections 2.3.1 and 2.3.2.1. The assay was incubated at 37 °C, the optimal temperature for all proteins, for 45 min, and can be observed in Figure 3-14.

![Activity assay gels](image)

*Figure 3-14. Activity assay gels for E. coli CSR603 with pQR1866 (exonuclease III), pQR1867 (λ exonuclease) and pQR1868 (T5 exonuclease). C is the control, EcoRV-digested λ DNA with nothing added but loading dye; 1, 3 and 5 are the non-induced protein extract assays for pQR1866, pQR1867 and pQR1868 respectively; and 2, 4 and 6 are the induced protein extract assays for pQR1866, pQR1867 and pQR1868 respectively.*

The lanes corresponding to the activity assays of induced cultures with the different exonuclease plasmids show significant degradation of the EcoRV-digested λ DNA compared to the non-induced in all cases. The slight degradation observed in the non-
induced samples is background digestion due to the nucleases normally present in the cell.

3.2.4 Effect of exonuclease overexpression on growth

Once all exonucleases were proven to be successfully cloned and expressed, one exonuclease combination had to be selected. For this, cultures of E. coli CSR603 harbouring pTTQ18 and each of the five different exonuclease plasmids were grown, including non-induced and induced cultures. Growth was tracked through absorbance measurements at 600 nm (referred to as OD\textsubscript{600}). Despite some E. coli cultures with certain plasmids grew slightly slower than others, induction time was set to 3 h in all for simplicity. Cultures were sampled over 24 h for analysis. The aims were to determine if there was any impact of exonuclease expression on growth, which exonuclease pair performs best in vitro and at which point after induction, and to quantify DNA content in cells as a measure of activity in vivo. Growth curves are depicted in Figure 3-15.
Figure 3-15. Growth curves of E. coli CSR603 harbouring the different exonuclease plasmids, non-induced (blue) and induced (red). 1 L flasks were grown at 37 °C and 250 rpm in duplicate for each condition over 24 h, and samples were taken hourly. The blue arrows indicate the point of induction, 3 h after inoculation. The error bars indicate the range between the minimum and maximum values.
Specific growth rates are not discussed herein due to the behaviour of induced pQR1868 and pQR1870 growth curves. Induced cultures tend to reach a lower cell density than their non-induced counterparts. This effect is especially prominent in the pQR1868 and pQR1870, where this difference is apparent just two hours post-induction (t = 5 h). Reduced cell growth upon induction in toxin expression has been previously reported (Larentis et al., 2014), and it would be expected to observe this when expressing DNA-damaging proteins in a strain with reduced DNA repair capabilities such as CSR603. In addition, in the SDS PAGE gels in section 3.2.3.1, it was also the induced pQR1868 samples where no obvious band corresponding to T5 exonuclease could be observed. This can be explained by T5 exonuclease being toxic to E. coli, so that a small amount of protein expressed immediately after induction exercises such selective pressure that growth stalls and the amount of exonuclease is too small to be observed on SDS PAGE. However, the cell population seems to recover sometime between 12 and 24 h for pQR1868, and around the 10 – 12 h mark for pQR1870; this may be due to outgrowth of cells that have lost the plasmid, which will have a greater advantage over induced plasmid-harbouring cells, or mutations developing that prevent further exonuclease expression.

When comparing cells expressing T5 exonuclease alone or together with exonuclease III, induced cultures with pQR1868 reach a final OD$_{600}$ = 1.5, whereas those with pQR1870 reach a final OD$_{600}$ = 3, matching its non-induced equivalent. It is possible that overexpression of two exonucleases, enabling DNA degradation from both 5' → 3' and 3' → 5' polarities, implies a much higher selection pressure that causes cells to develop strategies to circumvent toxic expression more rapidly than when just one exonuclease is expressed. Despite this quicker recovery, pQR1870 would be preferable to pQR1868 if an efficient genome ablation chassis needs exonucleases digesting in both polarities.

In terms of maximum cell density, there are some differences across the non-induced cultures, as displayed in Table 3-1. Under the given media and growth conditions and
for the control pTTQ18, the maximum $OD_{600} = 3.86$. With regards to the single-exonuclease plasmids, cultures harbouring pQR1866 and pQR1868 reach a similar value, whereas pQR1867 is 0.18 units below this maximum considering the standard deviation ($OD_{600} = 3.41$). This small difference is considered negligible and attributed to the inaccuracy of absorbance measurements via spectrophotometer: dry cell weights could have performed to track growth with higher accuracy, but absorbance measurements are quicker and deemed enough to be able to discern overall tendencies in growth. However, looking at the combined exonuclease plasmids, while pQR1869 displays a maximum growth within the range of the control ($OD_{600} = 3.72$), pQR1870 reaches a maximum $OD_{600} = 3.12$. This difference is too significant to be attributed to spectrophotometer inaccuracy.

Table 3-1. Maximum $OD_{600}$ reached by each non-induced culture of E. coli CSR603 with the different exonuclease plasmids.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>pTTQ18</th>
<th>pQR1866</th>
<th>pQR1867</th>
<th>pQR1868</th>
<th>pQR1869</th>
<th>pQR1870</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum $OD_{600}$</td>
<td>3.85 ± 0.23</td>
<td>3.6 ± 0.06</td>
<td>3.41 ± 0.03</td>
<td>4.08 ± 0.03</td>
<td>3.72 ± 0.06</td>
<td>3.12 ± 0.00</td>
</tr>
</tbody>
</table>

As a side note, the growth curve for CSR603 pQR1868 was done twice, and the second repeat is the one displayed in Figure 3-15 and analysed in this chapter. The comparison between both can be observed in Figure 3-16.
Figure 3-16. Side-to-side comparison of the E. coli CSR603 pQR1868 growth curve in Figure 3-15 (left) with the first attempt (right), non-induced (blue) and induced (red). The blue arrow indicates the point of induction, 3 h after inoculation. Both experiments were performed at 37 °C and 250 rpm, in duplicate and triplicate respectively. The error bars in the left graph indicate the minimum and maximum datapoints, and the standard deviation in the right graph.

In the first iteration, no difference was observed between non-induced and induced cultures; in fact, these were more similar to each other than the non-induced and induced cultures in pQR1866 and pQR1867, where a slight difference in cell density can be observed. The difference in maximum OD$_{600}$ is due to using 1 L flasks with larger baffles in the first run, which enabled a higher cell density (final OD$_{600}$ = 5). It is also worth mentioning that the growth curve for CSR603 pQR1870 took several attempts, since sometimes OD$_{600}$ would not decrease following induction and protein activity was not observed despite the presence of the correct plasmid was confirmed via sequencing. In those instances, the experiment was interrupted since exonuclease activity was not detected. This might be due to the variability in using complex media, LB broth, for these experiments: its exact composition is unknown due to the yeast extract and peptone it contains and introduces variability to the culture media (Sridhar and Steele-Mortimer, 2016). This will be discussed further in Chapter 5.
3.2.5 Visualisation of exonuclease-expressing *E. coli* CSR603 with a widefield microscope

As discussed in Chapter 1, section 1.5.3.4, extensive DNA damage triggers the SOS response, through which a series of events at a transcriptional level occur in order to help cells overcome this damage. Among other events, the SOS response induces SulA production, a cell division inhibitor in order to prevent cells transmitting damaged DNA to daughter cells (Huisman and D’ari, 1981). Consequently, cells elongate and form long filaments. Cell filamentation results in changes in the refractive index, which alters the relation between optical density measurements and cell density and in some cases cell number should be estimated via microscopy or viable count (Stevenson et al., 2016).

As seen in Figure 3-15, induction of T5 exonuclease (pQR1868) results in a lower OD$_{600}$, while this is no observed for exonuclease III (pQR1866) or λ exonuclease (pQR1867). It is possible that in the latter cases this increase in OD$_{600}$ does not correlate with a higher cell density and means that cells are filamenting instead due to the DNA damaged caused through exonuclease induction. For this reason, a visual comparison between exonuclease III and T5 exonuclease-expressing cells was performed: *E. coli* CSR603 harbouring plasmids pQR1866 or pQR1868 were grown in duplicates to exponential phase, at which point one culture of each was induced. Samples were harvested 4 h after induction, point at which OD$_{600}$ is still low in *E. coli* CSR603 pQR1868 samples in Figure 3-15, then fixed and visualised in a brightfield microscope. The images can be found in Figure 3-17 and Figure 3-18 for pQR1866 and pQR1868 respectively.
Figure 3-17. Brightfield microscope images of E. coli CSR603 pQR1866, non-induced (A) and induced (B). The samples were mounted on agarose pads and observed with an Olympus IX81 microscope with a 100x oil immersion objective. The white bars indicate 10 µm.
Figure 3-18. Brightfield microscope images of E. coli CSR603 pQR1868, non-induced (A) and induced (B). The samples were mounted on agarose pads and observed with an Olympus IX81 microscope with a 100x oil immersion objective. The white bars indicate 10 µm.
There is no elongation between non-induced and induced samples for either exonuclease III (pQR1866) or T5 exonuclease (pQR1868), although in the former case there is more filamentation in both non-induced and induced conditions. This indicates that OD$_{600}$ measurements are reliable for the purpose of growth comparison, and that exonuclease expression is not causing filamentation.

### 3.2.6 Effect of exonuclease expression on total DNA content

When performing the experiments to obtain the growth curves for the different exonuclease plasmids at both non-induced and induced conditions, 1 mL samples were taken at different points to extract total DNA and measure it via the PicoGreen assay in order to estimate the effect of exonuclease expression in vivo. This DNA quantification is portrayed in Figure 3-19 as raw data: after exonucleases induction, there will be a heterogeneous population in which some cells are expressing exonuclease that are degrading their DNA slowly; others in which this happens faster; and others that escape induction in which the DNA content does not change. For this reason, it is not accurate to normalise DNA content by cell density measured through OD$_{600}$. In addition, an attempt of normalisation of DNA content yielded contradictory results such as a lower DNA content prior to induction in some instances.
Figure 3.19. Total DNA quantification of the different E. coli CSR603 cultures expressing exonucleases including a control, expressed as mg of extracted DNA per mL of culture quantified via the PicoGreen assay. Measurements were taken in triplicate for each duplicate flask. Experimental conditions are identical to the growth curves: 2 L flasks at 37 °C and 250 rpm. The blue arrow indicates the time of induction, 3 h after inoculation. The error bar indicates the standard deviation. The sample at 3 h was taken just before IPTG addition.
There is a considerable increase in total DNA in the control pTTQ18 between 3 and 4 h, whereas there is not a great difference afterwards. Cells are likely to be transitioning to mid-exponential phase between those timepoints, since they will be more metabolically active and rapidly dividing. Exponentially growing *E. coli* contain several chromosomes per cell, and this number starts to decline before the transition to stationary phase occurs. Therefore, the ratio between DNA content and cell mass is not constant during exponential phase, and as growth progresses DNA content increases at a lower rate than cell density (Åkerlund et al., 1995). This explains why the amount of DNA isolated does not follow the same pattern as the growth curve, which displays the usual lag, exponential, stationary, and death phases. However, not all cultures had the same OD$_{600}$ at 3 h, which explains why this pattern is not observed in all graphs; in fact, this increase in DNA content between 3 and 4 h is only also observed in the non-induced pQR1868 culture (T5 exonuclease) and occurs later in the rest: between 4 and 5 h in pQR1866 (exonuclease III) and pQR1869 (exonuclease III and λ exonuclease), and between 5 h and 7 h in possibly pQR1867 (λ exonuclease; error bars make it difficult to determine) and especially pQR1870 (exonuclease III and T5 exonuclease). In the case of the latter, OD$_{600}$ upon induction was 0.57, whereas the rest were at OD$_{600} = 0.7 – 0.9$, which explains why the total DNA isolated from non-induced pQR1870 cultures is lower until 7 h. As mentioned in section 3.2.4, induction time was set at 3 h for all strains for simplicity.

As expected, total DNA content is lower in the induced cultures than in their non-induced counterparts. This difference is manifested at different times for different plasmids. Focusing on the single exonuclease plasmids, in the case of pQR1866 (exonuclease III) this happens 2 h post-induction (t = 5 h from beginning of growth); in pQR1867 (λ exonuclease), this takes 4 h (t = 7 h); and in pQR1868 (T5 exonuclease) it only takes 1 h (t = 4 h). The time window of the effect of exonuclease III is not reduced when expressed alongside λ exonuclease in pQR1869: this difference in DNA content is still only apparent
2 h after induction (t = 5 h), whereas it still just takes 1 h (t = 4) for T5 exonuclease expressed alongside exonuclease III in pQR1870 cultures.

The overall DNA content reduction is more pronounced in pQR1866 (exonuclease III) and pQR1868 (T5 exonuclease) than pQR1867, which suggests that λ exonuclease is the least efficient out of the three in terms of in vivo DNA degradation. The lowest DNA content due to expression of a single exonuclease is reached in the induced pQR1868 culture, which implies that T5 exonuclease is the most efficient nuclease. This also explains the slower cell growth post-induction discussed in section 3.2.4.

It is interesting that λ and T5 exonucleases have such different effects despite degrading DNA from the same polarity. However, despite both being exonucleases, these proteins are fundamentally different in terms of evolutionary origin, function and activity. T5 exonuclease displays sequence homology with the N-terminal domain of E. coli DNA polymerase I (Gutman and Minton, 1993), making it member of a protein family of exonucleases associated physically with a polymerase or in a complex including polymerases. In contrast, λ exonuclease belongs to a protein family not associated with such complexes. Upon hydrolysis, λ exonuclease releases dNMPs, dinucleotides and short oligonucleotides (Little, 1967a), while T5 exonuclease releases dNMPs up to 6-mer (Sayers and Ecksteins, 1990). λ exonuclease is also unable to hydrolyse nicked DNA (Carter and Radding, 1971) and can degrade ssDNA to a limited extent (Little, 1967), whereas T5 exonuclease can catalyse degradation on both nicked and ssDNA (Sayers and Ecksteins, 1990). In addition, the preferred substrate for λ exonuclease is 5'-phosphorylated DNA ends (Little, 1967), while T5 exonuclease will degrade dephosphorylated DNA. One main difference is that T5 exonuclease has an additional ‘flap’ endonuclease activity (Ceska et al., 1995). DNA flaps are DNA structures formed during DNA replication and repair, overhanging sections of ssDNA in which nucleotides are prevented from binding to their complementary base pair.
Considering this, T5 exonuclease seems preferable to λ exonuclease given its ability to initiate degradation from nicked DNA and ssDNA at a greater rate and not requiring a phosphate in the 5’ end of the polynucleotide chain. In terms of degradation products, T5 exonuclease generates smaller DNA products than λ exonuclease. It is therefore no surprise that expression of T5 exonuclease is more detrimental to cells than λ exonuclease; in fact, the first group who cloned and sequenced T5 exonuclease failed to obtain sufficient quantities to perform more comprehensive studies using pBR322 (Kaliman et al., 1986). Induced pQR1870 cultures (exonuclease III and T5 exonuclease) achieve the lowest DNA content, where this remains low up to 6 h post-induction ($t = 9$ h), when it starts increasing until it reaches a DNA content comparable to that of the non-induced culture. This mirrors what is observed in the growth curves (Figure 3-15), where cell density is seen to start increasing again at 9 h and reaches the same OD$_{600}$ value as the non-induced culture after overnight growth.

When comparing the two T5 exonuclease-coding plasmids, pQR1868 and pQR1870, the former would seem a better choice given the consistently lower DNA content and growth: expression of T5 exonuclease alone with pQR1868 together with the effect of background cellular nucleases may be enough. However, ultimately it was decided to select a combination of exonucleases digesting DNA from both polarities, so the choice still lies between pQR1869 and pQR1870. It would be expected that expression of exonuclease III added to λ exonuclease in pQR1869 would mean that the former could compensate one of the limitations of the phage exonuclease: exonuclease III is capable of initiating hydrolysis on both nicked dsDNA and ssDNA, but its capacity to degrade ssDNA is only 25 – 33% of the dsDNA degradation rate (Rogers and Weiss, 1980). It is a more versatile protein than λ or T5 exonuclease, having four different functions: 3’ → 5’ exonuclease, releasing dNMPs and ssDNA of the polarity it cannot digest (Richardson et al., 1964); it is an AP (apurinic) endonuclease, participating in base excision repair by cleaving the phosphodiester backbone of the AP site; it has RNaseH
activity, meaning it completely degrades the RNA strands from a DNA-RNA duplex; and it is a 3’-phosphatase (Richardson et al., 1964). Indeed, it does seem that expression of both exonuclease III and λ exonuclease in pQR1869 achieves a lower total DNA content than pQR1867 alone: ssDNA can now be somewhat digested because of exonuclease III, and hydrolysis of nicked DNA is possible.

This increased DNA degradation is not as obvious in pQR1870, where total DNA content is only marginally lower than in pQR1868. Both exonuclease III and T5 exonuclease can hydrolyze nicked DNA and T5 exonuclease can better digest ssDNA. Addition of exonuclease III implies that the 5’→3’ strand will be degraded down to dNMPs, while T5 exonuclease degrades down to 6-mers and λ exonuclease also releases short oligonucleotides. Given the characteristics of these exonucleases and the need for a 5’→3’ and a 3’→5’ exonuclease, pQR1870 seems a better choice for a genome ablation cassette if DSBs are introduced in the 3–9 h window, prior to cell growth recovery.

3.2.7 Loss of viability in cells expressing exonucleases

Seeing the effect of exonuclease overexpression on growth and DNA content, an additional experiment was performed in order to determine whether there is loss of viability associated with exonuclease induction. Strain CSR603 pQR1868 was selected due to the slow growth observed after induction. Cells were grown until early exponential phase, when a sample was taken and serial dilutions performed and plated on LB agar Amp or LB agar Amp IPTG. The colony forming units per mL (cfu / mL) were counted and results compared as seen in Table 3-2.
Table 3.2. Viable cell count of E. coli CSR603 pQR1868 in permissive (+Amp) or non-permissive (+Amp, IPTG) conditions. A cell sample was diluted and plated in triplicate on LB agar plates and incubated at 37°C until colonies were observed in order to determine the number of colony forming units per mL (cfu / mL).

<table>
<thead>
<tr>
<th></th>
<th>Amp</th>
<th>Amp + IPTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cfus/mL</td>
<td>$1.4 \cdot 10^7$ (± $4.4 \cdot 10^5$)</td>
<td>$1.4 \cdot 10^7$ (± $1.4 \cdot 10^6$)</td>
</tr>
</tbody>
</table>

The viable count was found to be identical for both permissive and non-permissive conditions, which seems to contradict the growth curves in Figure 3-15. This suggests that exonuclease expression causes a delay in growth, and that if liquid cultures were grow for long enough, induced cells would grow to similar levels to the non-induced cultures. Another possible explanation is that T5 exonuclease was not being expressed, like in the first attempt of the CSR603 pQR1868 growth curve in Figure 3-16, where cell growth did not stop post-induction due to an absence of exonuclease activity.

### 3.2.8 In vitro exonuclease activity assays

In order to compare the activity of the different exonucleases and their combinations, whole cell protein extracts were incubated with λ DNA digested with three different restriction enzymes to originate different ends: EcoRV, which creates blunt ends; Bpml, which cleaves leaving 3’ overhangs of 2 bp; and HindIII, which leaves 5’ overhangs of 4 bp. Samples were taken alongside those for total DNA content analysis, but in this case the volume taken was adjusted by OD$_{600}$ units (specifically, a volume equivalent to 1 mL of cells at an OD$_{600} = 0.5$). Methods of protein extract preparation and the activity assays are explained in sections 2.3.1 and 2.3.2.1. Only gels belonging to one of the induced replicates of each flask can be found in this section, in Figure 3-20 to Figure 3-34; the rest (the pTTQ18 control, non-induced cultures and the other replicate) can be found in the appendix (see section 8.5). Initially, the intention was to quantify DNA degradation through densitometric analysis; however, this method is too inaccurate to reflect small differences in degradation efficiency.
Figure 3-20. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-21. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 3-22. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-23. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 3-24. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-25. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 3-26. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-27. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on Bpml-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract. The 30 min assay at the 9 h timepoint should be discarded due to error in sample loading.
Figure 3-28. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-29. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 3-30. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-31. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 3-32. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 3-33. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Exonuclease activity can be immediately observed upon induction in all cases and for all substrates, since DNA degradation can be observed for the assays from samples taken just 1 h after induction (t = 4 h) after 30 min of incubation. As observed in the activity test gels on EcoRV-digested λ DNA (Figure 3-20, Figure 3-23, Figure 3-26, Figure 3-29 and Figure 3-32), all exonucleases and combinations are able to digest blunt-ended DNA, as well as DNA with 3' overhangs (BpmI-digested λ DNA gels: Figure 3-21, Figure 3-24, Figure 3-27, Figure 3-30 and Figure 3-33). However, it seems that DNA with 5' overhangs are more resistant to degradation (HindIII-digested λ DNA gels: Figure 3-22, Figure 3-25, Figure 3-28, Figure 3-31 and Figure 3-34). T5 exonuclease initiates digestion from blunt ends, or dsDNA with 5' and 3' overhangs (Garforth and Sayers, 1997); λ exonuclease can also digest blunt-ended DNA, dsDNA with 3' overhangs, and with 5' overhangs to a lesser extent (Mitsis and Kwagh, 1999); and exonuclease III can degrade dsDNA with blunt ends, 5' overhangs and 3' overhangs less than 4 nucleotides long (Hoheisel, 1993). Mitsis and Kwagh (1999) report a lower efficiency of λ exonuclease on 3' overhangs, substrate on which the enzyme has a Kcat/Km = 0.9 nM
as opposed to \(1.57 \text{nM}^{-1} \text{s}^{-1}\) on both blunt ends and 5' overhangs. No work in the literature has been found that compares \(K_{\text{cat}}/K_{\text{m}}\) for exonuclease III and T5 exonuclease under the same conditions for different substrates, but these results suggest these enzymes to also be more efficient at degrading blunt-ended DNA and with 5' overhangs than with 3' overhangs.

Interestingly, exonuclease activity is lost over time in both induced pQR1868 cultures (Figure 3-26, Figure 3-27 and Figure 3-28; and Figure 8-59, Figure 8-60 and Figure 8-61 in section 8.5 in the Appendix), both pQR1870 cultures (Figure 3-32, Figure 3-33 and Figure 3-34; and Figure 8-77, Figure 8-78 and Figure 8-79 in section 8.5 in the Appendix) and one of the pQR1869 replicates (Figure 8-68, Figure 8-69 and Figure 8-70 in section 8.5 in the Appendix). In pQR1870, this starts happening at the 9 - 10 h timepoint, which matches the gradual recovery in the growth curve starting at 9 h and in total DNA at 9 – 11 h. In the case of pQR1868 and the pQR1869 replicate, this seems to happen only at 24 h, even though these cultures did not behave differently in terms of growth or DNA content. This suggests that not only T5 exonuclease alone can exert an overwhelming selective pressure in the cell due to its toxicity, but that the combination of exonuclease III and \(\lambda\) exonuclease can have the same effect.

This is related to the marginally superior performance \textit{in vitro} of pQR1868 extracts (Figure 3-26, Figure 3-27 and Figure 3-28) than pQR1870 (Figure 3-32, Figure 3-33 and Figure 3-34); and of pQR1867 (Figure 3-23, Figure 3-24 and Figure 3-25) than pQR1869 (Figure 3-29, Figure 3-30 and Figure 3-31): in these cases, expression of two exonucleases simultaneously implies a larger selective pressure than expression of one alone. In addition, there is a greater exonuclease activity in the pQR1869 assays (Figure 3-29, Figure 3-30 and Figure 3-31) than in the pQR1870 ones (Figure 3-32, Figure 3-33 and Figure 3-34), and in pQR1867 than in pQR1868, despite T5 exonuclease being more efficient in terms of total DNA reduction. As explained at the beginning of this section, the culture volume taken for the protein extracts was adjusted by OD600 units, not by
amount of exonuclease present in the extract. A lower number of exonucleases produced in pQR1870 cultures due to the selective pressure exerted by T5 exonuclease expression would result in an equally lower abundance of exonucleases in induced pQR1870 extracts than in induced pQR1869 extracts, which would explain a higher efficiency of the latter in vitro: the small amount of T5 exonuclease produced just after induction is deleterious to cells, which will stop producing as much protein. This T5 exonuclease will degrade over time, so in vitro tests for later timepoints do not display much activity. In contrast, induced pQR1869 cells will continue expressing exonucleases, displaying in vitro activity even at later phases of induction. Therefore, even though in vitro tests were necessary to observe exonuclease activity on different substrates, which is much more difficult to observe in vivo, total DNA content is a better measure of exonuclease efficiency.
3.3 Conclusions

Three different exonucleases have been successfully cloned and expressed in *E. coli* CSR603. Cells harbouring the different exonuclease plasmids have been grown, monitoring cell growth, total DNA content and activity through *in vitro* assays. Exonuclease expression has been found to be toxic, but the DNA damage caused as a result does not induce elongation as observed in brightfield microscope imaging. T5 exonuclease is particularly toxic, and despite its expression is not obvious through SDS PAGE, its presence has been confirmed in *in vitro* activity assays. Cells tend to develop resistance against exonuclease expression as seen through loss of exonuclease activity overnight in T5 exonuclease-expressing cells, and in one instance when expressing both exonuclease III and λ exonuclease. When expressing exonuclease III alongside T5 exonuclease, cells recover and start growing again, and total DNA content increases. In addition, the identical viable count obtained when plating T5 exonuclease-encoding cells on LB Amp and LB Amp IPTG suggests that despite differences in non-induced and induced growth curves, cells tend to recover if grown for a sufficient period of time.

An exonuclease plasmid must be selected in order to move forward with the development of a genome ablation chassis. In terms of *in vivo* activity, the choice would lie between pQR1868 and pQR1870: expression of T5 exonuclease alone, not alongside an exonuclease digesting DNA from the 3′ → 5′ polarity, seems to be compensated by the presence of native exonucleases in the cell. However, pQR1870, encoding exonuclease III and T5 exonuclease, is still marginally better in terms of DNA content. *In vitro* assays reflect a greater exonuclease activity for pQR1869, although this is probably due to a lower quantity of exonucleases produced in pQR1870 cultures, highlighting that *in vitro* activity is not representative of performance *in vivo*. Therefore, the best approach seems to use pQR1870 and combining this with addition of endonucleolytic agent within the 3 – 12 h timeframe; in particular, at 4 h post-induction, where the lowest DNA content is achieved. This will be explored in the subsequent chapters.
Chapter 4. Use of DSB-inducing antibiotics with nuclease expression

4.1 Introduction

In Chapter 3, the combination of exonuclease III and T5 exonuclease was concluded to be most optimal for a genome ablation chassis. The other component of such a system, an efficient endonucleolytic agent, must be now found and assessed for efficiency in terms of reduction of viability and DNA degradation when combined with exonuclease expression. The initial objective of expressing different restriction enzymes was not fulfilled straightaway, and it took some time to achieve successful expression of any endonuclease. Therefore, the use of antibiotics to induce double-strand breaks (DSBs) was examined. Mitomycin C (MMC) and ciprofloxacin (CI) were selected as candidate drugs for this purpose. MMC is an antitumoral drug and CI is an antibiotic commonly used to treat bacterial infections. In vivo, MMC undergoes a chemical transformation and causes interstrand cross-links and same-strand links between guanines; when the nucleotide excision repair system tries to repair this, a DSB is formed (Roberts et al., 2014). In contrast, CI binds to subunit A of DNA gyrase as it unwinds the polynucleotide chain, stabilising the cleavage complex and generating a DSB (Snyder and Drlica, 1979).

In order to use these drugs as DSB-inducing agents, a range of concentrations was first tested, both individually and in combination by adding them to exponentially growing E. coli cultures. Exonuclease-expressing strain CSR603 was initially used, and when PvuIII was successfully cloned and expressed, this endonuclease was also co-expressed in this background. Strain DH5αZ1 was also tested, since it was the original strain in which the proof of concept for the expression system used for PvuIII was developed, as will be discussed in Chapter 5. In all cases, DNA was estimated via densitometry and cell viability was inspected via growth on LB agar.
4.2 Results

4.2.1 Evaluating different concentrations of MMC and CI

Different concentrations of MMC and CI were added to *E. coli* CSR603 in exponential phase of growth (OD\(_{600} = 1\)) and incubated at optimal growth conditions. OD\(_{600}\) was measured at time of addition and after an overnight incubation, and the difference calculated as a measure of comparison between conditions. 6 h after addition, 1 mL of culture was plated on LB agar, incubated overnight at 37 °C and examined for colony growth. All concentrations (detailed in Table 4-1) were tested in duplicate, and the results can be observed in Figure 4-1.

Table 4-1. Compilation of all combinations of MMC and CI assayed in Figure 4-1. Concentrations are specified for MMC and CI respectively. Abbreviations: MMC (mitomycin C), CI (ciprofloxacin).

<table>
<thead>
<tr>
<th>Sample</th>
<th>Concentration (µg/mL)</th>
<th>Sample</th>
<th>Concentration (µg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMC #1</td>
<td>1</td>
<td>MMC, CI #5</td>
<td>1, 200</td>
</tr>
<tr>
<td>MMC #2</td>
<td>5</td>
<td>MMC, CI #6</td>
<td>5, 1</td>
</tr>
<tr>
<td>MMC #3</td>
<td>10</td>
<td>MMC, CI #7</td>
<td>5, 10</td>
</tr>
<tr>
<td>CI #1</td>
<td>1</td>
<td>MMC, CI #8</td>
<td>5, 50</td>
</tr>
<tr>
<td>CI #2</td>
<td>10</td>
<td>MMC, CI #9</td>
<td>5, 100</td>
</tr>
<tr>
<td>CI #3</td>
<td>50</td>
<td>MMC, CI #10</td>
<td>5, 200</td>
</tr>
<tr>
<td>CI #4</td>
<td>100</td>
<td>MMC, CI #11</td>
<td>10, 1</td>
</tr>
<tr>
<td>CI #5</td>
<td>200</td>
<td>MMC, CI #12</td>
<td>10, 10</td>
</tr>
<tr>
<td>MMC, CI #1</td>
<td>1,1</td>
<td>MMC, CI #13</td>
<td>10, 50</td>
</tr>
<tr>
<td>MMC, CI #2</td>
<td>1, 10</td>
<td>MMC, CI #14</td>
<td>10, 100</td>
</tr>
<tr>
<td>MMC, CI #3</td>
<td>1, 50</td>
<td>MMC, CI #15</td>
<td>10, 200</td>
</tr>
<tr>
<td>MMC, CI #4</td>
<td>1, 100</td>
<td>Control</td>
<td>-</td>
</tr>
</tbody>
</table>
Figure 4-1. Change in absorbance at 600 nm post-drug addition in E. coli CSR603 pQR1870 expressed as the difference between the absorbance of the culture at $t = 24$ h and at $t = 3.5$ h (time of antibiotic addition). The combinations of MMC and CI detailed in Table 4-1 were added in duplicates to 10 mL E. coli cultures 3.5 h after inoculation, grown at 37 °C and 250 rpm, and absorbance at 600 nm was measured after 24 h. At that point, 0.5 mL culture was resuspended in 100 µL LB and plated on LB agar plates. The error bars indicate the range between the minimum and maximum values of absorbance at 600 nm. Asterisks (*) indicate the conditions in which there were survivors.

Abbreviations: MMC (mitomycin C), CI (ciprofloxacin).
There is an interesting juxtaposition in the effects of MMC and CI: while CI causes a significant reduction in $OD_{600}$ starting from the third highest concentration, it does not imply loss of viability at the tested concentrations, and while this reduction in $OD_{600}$ is not as clear when adding MMC, even at the lowest concentration assayed no survivors grew on LB agar.

Addition of MMC at its lowest concentration (MMC #1, 1 µg / mL) guarantees no survivors, so this concentration was selected. Moreover, CI at three different concentrations was added: 10 µg / mL (CI #1), which allows for cell growth manifested via an $OD_{600}$ increase; and 100 (CI #4) and 200 µg / mL (CI #5), which cause a decrease in $OD_{600}$. The selected combinations are therefore MMC+CI#1, MMC+CI#4, and MMC+CI#5.

4.2.2 Combining drug addition with exonuclease overexpression in *E. coli* CSR603 pQR1870

Once the MMC and CI concentrations were selected, the next experiment entailed growing *E. coli* CSR603 pQR1870, inducing exonuclease expression at 3 h and adding DSB-inducing antibiotics at 7 h. Samples were taken at 10, 13 and 24 h. At all these timepoints, cell growth was monitored by $OD_{600}$ and 1 mL samples were taken for DNA extraction and quantification. Cell viability was assessed by taking a 0.5 mL sample at 24 h and plating it on LB agar.

As seen in the growth curve in Figure 4-2, cell density is visibly lower in the induced samples than the non-induced, which proves that exonucleases are being expressed. This is important to denote, since inconsistencies were observed repeatedly in $OD_{600}$ upon induction as mentioned in Chapter 3, section 3.2.4, where growth curves for CSR603 pQR1870 took several attempts and two growth curves for CSR603 pQR1868 were different (Figure 3-16). In some cases, cell density did not decrease post-induction, and cells were discarded and the experiment repeated.
Figure 4-2. Growth curves of E. coli CSR603 pQR1870, either non-induced (n/I) or induced (Ind.), at different MMC and CI concentrations. Cultures were grown in duplicate for each condition in 250 mL flasks at 37 °C and 250 rpm, and if pertinent, induced after 3 h. Antibiotics were then added at 7 h. The error bars indicate the range between the minimum and maximum values of absorbance at 600 nm. n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin. ‘no MMC, CI’ means no antibiotics were added; ‘low / med / high MMC, CI’ means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.

The effect of antibiotic addition in non-induced cells is more obvious than in induced cultures, at least for the low and the medium concentrations: since cell growth has already been impaired by exonuclease expression in the induced samples, the effect of further DNA damage is not as noticeable as in exponentially growing cells. In all cases, cell growth on LB agar was only observed in those cells that had no MMC or CI added, and no viable cells were observed from plating of cultures with induced exonucleases treated with MMC and CI at 7 h. One replicate for the non-induced and the induced controls with no antibiotic addition were contaminated after the overnight incubation, so this data was discarded. This contamination was identified as a different colony colour (yellow instead of white), morphology and smell than that of E. coli.

Interestingly, in this case, induced cells with no added antibiotic do not see recovery in their growth, instead reaching OD₆₀₀ levels similar to the non-induced cultures. This contradicts the CSR603 pQR1870 growth curves in Chapter 3 (see Figure 3-15), in which induced cultures reached the same cell density levels as the non-induced counterparts after overnight growth. It is the only instance throughout this work that no recovery in
growth in CSR603 pQR1870 cells was observed. There is no clear explanation as to why this happened.

1 mL samples were taken at the same time as OD_{600} was read, and total DNA was extracted, run in agarose gels and analysed via densitometric analysis on ImageJ. This densitometric analysis was performed to quantify the variation in DNA content from the time of induction, 3 h, to each of the timepoints. The graph depicting this variation in DNA content with respect to the time of induction is depicted in Figure 4-3. The DNA gel used for this analysis is depicted in Figure 4-4.

![Figure 4-3](image-url)

**Figure 4-3.** Densitometric analysis of total DNA content in E. coli CSR603 pQR1870 with or without exonuclease expression and at different MMC, CI concentrations. Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h, in the same agarose gel. Cells were induced at 3 h after inoculation; antibiotics were added at 7 h. The error bars indicate the range between the minimum and maximum values.

n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin. 'no MMC, CI' means no antibiotics were added; 'low / med / high MMC, CI' means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.
Some bands were overexposed, but it was deemed unnecessary to repeat this given that the goal was to find a combination that yielded no visible DNA in the gel; therefore, overexposure was preferable to underexposure, which would have caused the faintest bands to disappear. In addition, the data was not normalised by OD$_{600}$ given that this would have magnified the inaccuracies.

As expected, DNA content is overall lower in the induced samples than the non-induced, which is particularly noticeable at $t = 7$ h. DNA content increases by at least 134% in the
non-induced samples, whereas induced samples are estimated to decrease to 86 % the DNA content at 3 h. MMC and CI addition occurs at \( t = 7 \) h, and three hours later at \( t = 10 \) h the effect is noticeable, with the different DNA contents decreasing on par with the increasing concentrations for both subsets. The lowest DNA concentration achieved is approximately 5 % that of the content at 3 h, and is achieved in the induced sample after 24 h at the highest antibiotic concentrations. At 13 h, the medium and highest concentrations achieve similar results (7 % and 6 % approximately).

Unfortunately, the goal of finding an antibiotic combination that yielded no DNA was not achieved and other avenues had to be explored.

4.2.3 Combining MMC and CI with endo- and exonuclease expression

Later in this work, several endonucleases were cloned and expressed in \( E. coli \) DH5αZ1 (to be discussed in Chapter 5). It was decided to co-express restriction enzyme \( PvuII \) and the exonucleases, combining this with MMC and CI addition, in order to achieve genome ablation. First, this system was tested in \( E. coli \) CSR603, and then in DH5αZ1 since this was the model organism originally used to develop the expression strategy.

In this case, as will be explained in Chapter 5, plasmids pQR1871 and pQR1872 encoding exonucleases and \( PvuII \) are used instead of pQR1870.

4.2.1.1 CSR603

As in the previous experiment, \( E. coli \) (harbouring plasmids pQR1871 and pQR1872 instead of pQR1870) was induced at 3 h and DSB-inducing antibiotics were added at 7 h. Samples were taken at 10, 13 and 24 h. At all these timepoints, cell growth was monitored by OD\(_{600}\) and 1 mL samples were taken for DNA extraction and quantification. Cell viability was assessed by taking a 0.5 mL sample at 24 h and plating it on LB agar.

In this case, to simplify the experiment, one large flask was grown until time of addition, point at which it was split; therefore, there was only one sample for the non-induced and the induced subsets, and only one DNA extraction was performed for both subsets at
3 h and 7 h. Like in the previous experiment, growth curves were performed (as seen in Figure 4-5) and densitometric analysis done (Figure 4-6; the DNA gel used for these calculations can be found in the Appendix, Figure 8-80).

![Growth curves](image)

**Figure 4-5.** Growth curves of E. coli CSR603 pQR1871 pQR1872, either non-induced (n/I) or induced (Ind.), at different MMC and CI concentrations. Cultures were grown in duplicate for each condition in 250 mL flasks at 37 °C and 250 rpm, and if pertinent, induced after 3 h. Antibiotics were then added at 7 h.

The error bars indicate the range between the minimum and maximum values of absorbance at 600 nm. n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin. ‘no MMC, CI’ means no antibiotics were added; ‘low / med / high MMC, CI’ means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.

No cells grew on LB agar as part of the viability test when antibiotic was added. The effect of antibiotic addition is similar to the growth curves in Figure 4-3, although in this case antibiotic concentration makes no difference in the induced samples, where all three growth curves are very similar: since cells are already burdened with expression of an endonuclease, the presence of additional DSB-inducing agents must be sufficient.

In this case OD₆₀₀ does increase by the 13 h timepoint and after overnight growth it has reached the same level as the non-induced cultures.
Figure 4-6. Densitometric analysis of total DNA content in E. coli CSR603 pQR1871 pQR1872 with or without exonuclease expression and at different MMC, CI concentrations. Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h, in the same agarose gel. Cells were induced at 3 h after inoculation; antibiotics were added at 7 h. The error bars indicate the range between the minimum and maximum values.

n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin.

‘no MMC, CI’ means no antibiotics were added; ‘low / med / high MMC, CI’ means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.

The DNA gel used for these calculations can be found in the Appendix, Figure 8-80.

In this case, for the induced samples, the decrease in DNA content is greater than when expressing exonuclease only (40% versus 32%). The effect of antibiotic addition is similar, and under induced and maximum antibiotic addition conditions, the DNA content after 24 h is 7%, close to the previously obtained 5%. Nevertheless, the goal of finding a set of conditions at which DNA content is reduced to zero is not reached here either, despite again all cells are killed, as demonstrated by absence of growth on LB agar upon antibiotic addition.

4.2.1.2 DH5αZ1

The decision to test this strain for DNA reduction with restriction enzyme and exonuclease expression was made because this is the strain in which the proof of concept of the expression system for the endonuclease was developed. First, the same experiment to determine optimal MMC and CI concentrations was performed as
explained in section 4.2.1, but with *E. coli* DH5αZ1 pQR1870, and the results are depicted in Figure 4-7.
Figure 4-7. Change in absorbance at 600 nm post-drug addition in *E. coli* DH5αZ1 pQR1871 pQR1872 expressed as the difference between the absorbance of the culture at \( t = 24 \) h and at \( t = 3 \) h (time of antibiotic addition). The combinations of MMC and CI detailed in Table 4-1 were added in triplicates to 10 mL *E. coli* cultures 3 h after inoculation, grown at 37 °C and 250 rpm, and absorbance at 600 nm was measured after 24 h. At that point, 0.5 mL culture was resuspended in 100 μL LB and plated on LB agar plates. The error bars indicate the range between the minimum and maximum values of absorbance at 600 nm. Asterisks (*) indicate the conditions in which there were survivors.

Abbreviations: MMC (mitomycin C), CI (ciprofloxacin).
The pattern is very similar to Figure 4-1, with CSR603 pQR1870. A difference between both strains is survival after addition of MMC and CI separately. Addition of the lowest MMC concentration inhibited any CSR603 growth on LB agar, while one lone colony grew in one of the two replicates of DH5αZ1 at 1 µg / mL MMC. The survival of this colony could mean that CSR603 is marginally more sensitive to MMC than DH5αZ1, which is explained through their different genotypes: CSR603 is *phr uvrA recA*, meaning it is not capable of DNA repair, while DH5αZ1 is just *recA* and can maintain repair pathways not involving recombination.

On the other hand, DH5αZ1 only survived the lowest concentration of CI. This can be due to their different genotypes: in the 1970s and 80 many nalidixic acid-resistant mutants of *E. coli* were isolated to better study DNA gyrase (Inoue et al., 1978) (Yamagishi et al., 1981), and both strains derive from different mutants; CSR603 is *gyrA98*, and DH5αZ1 is *gyrA96*, meaning that both are resistant to nalidixic acid. CI is a second-generation fluoroquinolone, and nalidixic acid is a first-generation quinolone. The fact that the mode of action of both is by binding to subunit A of DNA gyrase but both strains are susceptible to CI suggests that both quinolones are slightly different; for instance, CI might bind more strongly to gyrase than nalidixic acid. In fact, nalidixic acid activity is blocked by addition of protein synthesis inhibitors like chloramphenicol, implying that protein synthesis is necessary for quinolone action (Deitz et al., 1966); however, fluoroquinolone sensitivity persists in presence of chloramphenicol (Lewin and Amyes, 1990).

Despite the higher resistance of DH5αZ1 to MMC, the lowest concentration was still selected since it was assumed that the synergistic effect of MMC and CI should be sufficient to guarantee no survival, especially given the increased susceptibility to CI. The same experiment was then performed; growth curves are depicted in Figure 4-8, the densitometric analysis in Figure 4-9, and the DNA gel in Figure 8-81 in the Appendix, section 8.6.
Figure 4-8. Growth curves of E. coli DH5αZ1 pQR1871 pQR1872, either non-induced (n/I) or induced (Ind.), at different MMC and CI concentrations. Cultures were grown in duplicate for each condition in 250 mL flasks at 37 °C and 250 rpm, and if pertinent, induced after 3 h. Antibiotics were then added at 7 h.

The error bars indicate the range between the minimum and maximum values of absorbance at 600 nm. n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin. 'no MMC, CI' means no antibiotics were added; 'low / med / high MMC, CI' means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.

As in the previous two experiments, there was only colony growth on plates with no MMC or CI. Whereas antibiotic addition does have a clear effect on cell growth in the non-induced cultures, this is not as obvious in the induced samples. This can be explained like in the previous case: endonuclease expression is already causing a significant burden on the cell, so additional DSB-inducing antibiotics do not have further impact on growth. Additionally, cell growth does not recover like it was observed in Chapter 3 with pQR1870 and in Figure 4-5.
Figure 4-9. Densitometric analysis of total DNA content in E. coli DH5αZ1 pQR1871 pQR1872 with or without exonuclease expression and at different MMC, CI concentrations. Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h, in the same agarose gel. Cells were induced at 3 h after inoculation; antibiotics were added at 7 h. The error bars indicate the range between the minimum and maximum values. 

n/I stands for non-induced; Ind., for induced; MMC, mitomycin C; CI, ciprofloxacin.

‘no MMC, CI’ means no antibiotics were added; ‘low / med / high MMC, CI’ means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively.

The DNA gel used for these calculations can be found in the Appendix, Figure 8-81.

While OD_{600} remains at a similar level, there is a difference in DNA content between the non-induced and the induced samples, regardless of the large error bars in the former. The lowest DNA achieved is in the induced culture at the highest antibiotic concentration. The medium concentration sample of this subset seems to yield a slightly higher DNA content than the highest, but this is considered to be an artefact due to the use of densitometry as method for analysis. As occurred previously, in those samples to which antibiotics had been added, no colonies grew on LB agar.

### 4.2.4 Comparison of all strategies

Figure 4-10 summarises the densitometric analysis of the induced samples in all three experiments for ease of comparison.
Figure 4-10. Comparison between the different induced samples for the three main experiments in this chapter. The data is the same as in Figure 4-3, Figure 4-6 and Figure 4-9. Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h, in the same agarose gel. Cells were induced at 3 h after inoculation; antibiotics were added at 7 h.

n/a stands for no antibiotic addition; MMC, mitomycin C; CI, ciprofloxacin; RE: restriction enzyme (PvuII).

‘no MMC, CI’ means no antibiotics were added; ‘low / med / high MMC, CI’ means addition of MMC at 1 µg / mL and CI at 10, 100 and 200 µg / mL respectively. The error bars indicate the range between the minimum and maximum values.
In all cases there is a reduction in DNA content upon induction, which is more pronounced when *PvuII* accompanies exonuclease expression. In addition, there is not a large difference from 10 h to 24 h for all three conditions and for the different concentrations. The one difference is the non-addition control for CSR603 + restriction enzyme + exonucleases, which will be discussed in Chapter 5. Moreover, when comparing the highest MMC and CI concentrations amongst the three different strains, DH5αZ1 with restriction enzyme and exonuclease expression is the least efficient, since DNA falls to 20-25 % while in CSR603 this decrease to 5 – 10 %.
4.3 Discussion and conclusions

The goal of this chapter was to find a way to decrease DNA content and viability to zero via protein expression combined with DSB-inducing antibiotics. Cell viability has been found to be successfully prevented as expressed by lack of viable colony growth on LB agar, as long as MMC is added at 1 µg / mL alongside CI. *E. coli* strains CSR603 and DH5αZ1 have been found to have different tolerances towards CI, the latter being more sensitive to the fluoroquinolone despite having a higher capability of DNA repair. This difference is attributed to be due to their *gyrA* mutations occurring in different codons. Moreover, cultures expressing *Pvu*II and exonucleases do not display major changes in growth upon antibiotic addition because they are already burdened by restriction enzyme expression. Addition of MMC at 1 µg / mL and CI at 200 µg / mL to CSR603 expressing exonucleases with or without *Pvu*II is the most efficient combination for DNA reduction.

This strategy ensures no cell viability, which is still not possible when employing nuclease kill switches. In addition, most works involving nuclease kill switches to degrade intracellular DNA and trigger cell death only analyse the impact on cell viability, not including an examination of DNA content (Elešnik et al., 2016; Gallagher et al., 2015; Li and Wu, 2009). Exceptions are inclusion of a comet assay to show only a qualitative degradation of chromosomal DNA (Balan and Schenberg, 2005); use of qPCR targeting the specific loci targeted by a CRISPR device to prove the expected site is cleaved (Caliando and Voigt, 2015); and DNA degradation through nuclease expression shown via increase of radiolabelled acid-soluble material (Ahrenholtz et al., 1994). This work aims to prove there is both DNA reduction and a reduction in viability and to evaluate this strategy for construction of a DNA-free chassis.

As seen in this chapter, addition of DSB-inducing antibiotics is not enough for development of a genome ablation methodology: there is DNA left under all conditions explored in this chapter. However, the quality of this DNA may prevent horizontal transfer, since it might be too damaged to be acquired via natural transformation and
assimilated by other organisms or, in the case that this chassis acquired exogenous DNA, to allow for recombination. In addition, this defect in quality may prevent the chassis from thriving if released to the environment. Early studies on the effect of MMC on DNA showed that DNA extracted from MMC-treated \textit{B. subtilis} and \textit{E. coli} displayed a much lower transformation efficiency (Iyer and Szybalski, 1963).

Regardless, a genome ablation strategy that guarantees complete removal of DNA is still desirable. Alternative restriction enzymes to \textit{PvuII}, or other DSB-inducing proteins, may be more efficient. This will be discussed in the next chapter.
5 Chapter 5. DNA-damaging protein expression as a DSB-inducing agent

5.1 Introduction

In Chapter 4, the effect of addition of DSB-inducing antibiotics MMC and CI combined with exonuclease expression was studied. Whereas antibiotic addition ensures no survival, it is not optimal in terms of DNA reduction: there is DNA remaining even when expressing exonucleases at the highest antibiotic concentrations. Moreover, addition of antibiotics to a cell culture becomes significantly expensive as the culture is scaled up and it should be avoided whenever possible. The possibility of inducing DSB via protein expression is therefore a more attractive option.

There are many types of endonucleases in nature, as mentioned in Chapter 1, section 1.5.2.1. Examples include restriction enzymes, which are endonucleases that cleave DNA as a means to defend the cell against bacteriophages; bacteriophage endonucleases, which are employed by the virus to degrade host DNA and scavenge nucleotides for its own DNA replication; and colicins, toxins secreted by bacteria to counter competition by other organisms in the environment, some of which have DNAse activity. Other proteins, such as gyrase poisons, induce DNA cleavage indirectly by binding to DNA gyrase as it unwinds the polynucleotide chain, stabilising the cleavage complex and resulting in a DSB.

DNA-cleaving proteins have been used previously in kill switches (Ahrenholtz et al., 1994; Caliando and Voigt, 2015; Elešnik et al., 2016; Torres et al., 2003), but these rely on the nucleases as a suicide mechanism and do not consider them as tools for a DNA-free chassis as is done in this work. In this chapter, four different DSB-inducing proteins are cloned and expressed in *E. coli* DH5αZ1: restriction enzymes *Pvu*II and *Hpa*II, DNA gyrase poison ccdB, and viral endonuclease denA. These proteins were selected based on their non-patented status, sequence confirmed, and activity characterised. These are
expressed alongside exonucleases and evaluated in terms of their impact reduction of DNA content and killing efficiency via decrease in viability counts. The difficulties in expressing potentially lethal proteins is also addressed: out of the four different expression methods used, only one involving regulation via recombinase induction was successful. In addition, the influence of the time of nuclease induction on these two metrics is examined through DNA reduction and escape frequency, a metric often cited when discussing kill switches: the minimally acceptable escape frequency is $10^{-8}$ (National Institutes of Health, 2016). Finally, the increase in cell viability observed over time after nuclease induction is also addressed.
5.2 Results and discussion

5.2.1 Unsuccessful attempts at cloning DSB-inducing proteins

Several attempts to clone DSB-inducing proteins failed throughout this project, highlighting the difficulties entailing expression of deliberately toxic proteins. Several approaches were considered, ranging from cloning in high-copy to low-medium copy plasmids, and first expressing a methyltransferase and then the cognate restriction enzyme.

5.2.1.1 Cloning under the tac promoter in pTTQ18 and pQR445

In one first attempt, restriction enzymes *Pvu*II, *Bam*HI, *Not*I and *Hpa*II were PCR-amplified from their respective genomic DNA to be cloned and expressed in pTTQ18 via CPEC in the same manner as in Chapter 3. Details on the cloning procedure can be found in Chapter 2, section 2.2.6.3 and a detailed diagram in Chapter 3, Figure 3-5. Briefly, the inserts were amplified to include 21 - 25 bp homology with pTTQ18, a stop codon to interrupt the *lacZα* cistron, An RBS and the restriction enzyme CDS. 10 µL of the PCR reaction were loaded in agarose gels, which are displayed in Figure 5-1, Figure 5-2 and Figure 5-3.

![Image](image-url)

*Figure 5-1. PCR amplification of the NotI (1167 bp; lanes 1-3, blue arrow) and HpaII (1077 bp; lanes 4-6) genes (green arrow), for cloning via CPEC into pTTQ18. The bands at 100 bp are primer dimers (orange square). The template DNA was N. otiditis-caviarum and H. influenzae genomic DNA respectively. 5 µL of the PCR reaction were loaded. The ladder (M) is 100 bp ladder (NEB).*
Figure 5-2. PCR amplification of the BamHI gene (657 bp; lanes 1-2, blue arrow) for cloning via CPEC into pTTQ18. The bands at 100 bp are primer dimers (orange square). The template DNA was B. amyloliquefaciens genomic DNA. 5 µL of the PCR reaction were loaded. The ladder (M) is 100 bp ladder (NEB).

Figure 5-3. PCR amplification of the PvuII gene (474 bp; lanes 1-2, blue arrow) for cloning via CPEC into pTTQ18. The bands at 100 bp are primer dimers (orange square). The template DNA was the pPVU1 plasmid from P. vulgaris. 5 µL of the PCR reaction were loaded. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).

When attempting to clone HpaII into pTTQ18, not a single positive colony was obtained. For the other enzymes, plasmid DNA isolated from two colonies at least were sent to sequencing using primers #M13 rev (-49) and #M13 uni (-21) supplied by Eurofins (sequence available in the Appendix, section 8.2.1). Nevertheless, few colonies were obtained, and none were found to have the correct sequence. In the case of NotI, out of the only two positive colonies obtained, one only contained the first 18 nt and final 19 nt of the CDS, missing the 1115 nt in between; the other one had multiple frame-shifting deletions in the sequence. When attempting to clone PvuII, the four colonies obtained and sent for sequencing were found to have different minor deletions: a 4 and a 5 nt deletion at the beginning of the CDS (nts 3-6 and 10-14 respectively), single nt deletions at positions 11 and 20, and a 4 nt deletion close to the end of the CDS (nts 455-458). Finally, in the case of BamHI, six positive colonies were obtained with different mutations:
one with a single nt deletion (nt 18), one with two 1-nit deletions (positions 28 and 29),
one with a 1 nt insertion at position 34, one with a 1 nt insertion at position 19, one with
a C→T mutation resulting in a residue change (A139D), and one with an A→T mutation
resulting in an E111V. This particular mutation has been obtained in at least one other
work: E111 has been found to have a role in catalysis and a E111-deficient strain binds
to DNA but does not cleave it (Dorner and Schildkraut, 1994). While A139 has not been
described to have a role, the fact that this mutation was also obtained may imply that it
also participates in the mode of action of BamHI.

Given the failure to clone these proteins under the tac promoter in pTTQ18, cloning was
attempted in a plasmid with a lower copy number. Vector pQR445 was selected, a low-
copy vector with an RSF1010 origin of replication, for cloning of the PvuII CDS into the
BamHI site under control of the tac promoter via CPEC. The plasmid map of pQR445
can be found in the Appendix, section 8.1.2. Details on the cloning procedure can be
found in Chapter 2, section 2.2.6.4. A PCR was performed to amplify the PvuII CDS and
introduce homology with the pQR445 vector at either side of the BamHI site, and a small
volume of the PCR reaction was run in an agarose gel as seen in Figure 5-4.

![Figure 5-4. PCR amplification of the PvuII CDS (blue square; 546 bp fragment in lane 1) for
cloning via CPEC in pQR445. The template DNA is the pPVU1 plasmid from P. vulgaris. 10 µL
of the PCR reaction were loaded. The ladder (M) is the 100 bp ladder (NEB).](image-url)
5 µL of the CPEC reaction were transformed, but no colonies were obtained. Cloning was also attempted in vectors containing a \textit{araBAD} and \textit{rha} promoters, which are less leaky than the \textit{tac} promoter, but no colonies were obtained either. Therefore, a different approach was considered: using methyltransferases prior to cloning the restriction enzyme.

5.2.1.2 \textbf{Using methyltransferases: pACYC184 and pPVU1t}

This approach involves first obtaining a strain in which the cognate methyltransferase for the restriction enzyme was active: a methyltransferase would be constitutively expressed in the cell in a lower copy plasmid than the restriction enzyme, so that upon induction, expression of the restriction enzyme would be overwhelming for the cell to an extent that the DNA would not be methylated quickly enough. Expression of restriction enzymes in strains expressing their cognate methyltransferase is common when overexpressing restriction enzymes for later purification (Bao et al., 2008; Jack et al., 1991).

The methyltransferase genes for \textit{NotI}, \textit{PvuII}, \textit{BamHI} and \textit{Hpal} were PCR-amplified and cloned into pACYC184 under the chloramphenicol resistance promoter (CAT), which would enable constitutive expression. The plasmid map for pACYC184 can be found in the Appendix, section 8.1.3, and the maps of the methyltransferase plasmids in sections 8.1.9 to 8.1.12. Details of the cloning procedure can be found in Chapter 2, section 2.2.6.5. The gels of the PCR amplifications of the methyltransferase genes can be found in Figure 5-5, Figure 5-6 and Figure 5-7.
Figure 5-5. PCR amplification of the NotI (1127 bp; lanes 1 and 2, blue arrows) and the PvuII methyltransferase genes (1076 bp; lanes 3 and 4, green arrows). The band at 8 kb (orange squares) corresponds to template plasmid DNA containing the methyltransferase sequences (not mentioned in this work). 5 µL of the PCR reaction were loaded. The ladder (M) is 1 kb DNA Ladder (NEB).

Figure 5-6. PCR amplification of the BamHI methyltransferase gene (1337 bp; lane 1, blue arrow). The band at 5 kb (orange square) corresponds to template plasmid DNA containing the methyltransferase sequence (not mentioned in this work). 5 µL of the PCR reaction were loaded. The ladder (M) is 1 kb DNA Ladder (NEB).
Figure 5-7. PCR amplification of the HpaII methyltransferase gene (1142 bp; lanes 1 and 2, blue arrows). The band at 8 kb (orange squares) corresponds to template plasmid DNA. 5 µL of the PCR reaction were loaded. The template DNA is plasmid DNA containing the methyltransferase sequence (not mentioned in this work). The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).

Since no colonies were obtained in the vector-only transformation control, colonies were picked and plasmid DNA was sent to sequencing directly with oligonucleotides #S-01 and #S-02 (sequences available in the Appendix, section 8.2.1). The sequence of all methyltransferases was found to be correct. These plasmids were cloned using E. coli HB101. This strain has a hsdS20 (rB−, mB−) genotype, meaning it is deficient in the endonuclease and methylase of the E. coli EcoKI restriction-modification system. DNA propagated in these strains will not be methylated, and DNA transformed into this strain will not be cleaved. HB101 is the strain classically used when cloning exogenous restriction-modification systems (Brooks et al., 1989).

In order to test that these methyltransferases worked in vivo, genomic DNA plugs of different E. coli HB101 strains were prepared and challenged with their cognate restriction enzymes. The procedure for genomic DNA plug preparation and FIGE can be found in Chapter 2, section 2.2.1.12. As controls, individual plugs of E. coli HB101 with no methyltransferases were digested with NotI, PvuII, BamHI or HpaII to verify that the enzymes can degrade genomic DNA in agarose plugs. Then, genomic plugs of HB101 with the different methyltransferase plasmids were prepared and digested with their cognate restriction enzyme. The resulting gel can be observed in Figure 5-8.
No methyltransferase activity is observed: for all methyltransferase-harbouring strains, genomic DNA is digested in the same manner as the strains with no plasmids. No explanation for this could be found, given that the cloning was successful and the sequence was verified.

Given the failure to clone and express successfully these methyltransferases using pACYC184, it was decided to use a plasmid with the original restriction-modification system: the pPVU1 plasmid from *P. vulgaris*, which codes for the *Pvu*II restriction enzyme and methyltransferase. More details on the cloning procedure can be found in Chapter 2, section 2.2.6.6. Briefly, to enable manipulation of the plasmid in *E. coli*, a tet\(^R\) cassette was first introduced via ligation in the pPVU1 site, obtaining plasmid pPVU1t. Then, the restriction enzyme gene was deleted and a *Pvu*II site was introduced instead.
to ensure that the plasmid is also cleaved upon triggering genome ablation, yielding plasmid pPVU1t2. The plasmid maps for pPVU1, pPVU1t and pPVU1t2 can be found in the Appendix, sections 8.1.13 to 8.1.15. A gel with linearised pPVU1 and the tetR cassette is depicted in Figure 5-9.

![Gel with linearised pPVU1 and the tetR cassette](image)

*Figure 5-9. Gel with 50 ng of the EcoRI-digested pPVU1 plasmid (4.67 kb; blue arrow in lane 1) and 100 ng of the tetR cassette (1.3 kb; green arrow in lane 2). The ladder (M) is 1 kb DNA Ladder (NEB).*

After ligation, one tetracycline-resistant colony grew harbouring pPVU1t, which was then used as template to PCR amplify it excluding the PvuI gene and introducing the PvuII site. The PCR product, depicted in Figure 5-10, was then digested with PvuII and ligated overnight prior to transformation into *E. coli* HB101.

![PCR amplification of pPVU1t for linearization](image)

*Figure 5-10. PCR amplification of pPVU1t for linearization. Lane 1 is 1 µL of the PCR reaction; lane 2 is 20 µL (5.5 kb; blue arrows). The fragment at 100 bp (orange square) is a primer dimer. The fragment was gel-purified. The ladder (M) is GeneRuler DNA Ladder Mix (NEB).*

48 colonies were obtained, and none in the vector-only control. One colony was picked and the assembled product was confirmed via digestion with BamHI and PvuII to verify
that both were able to linearise the plasmid, meaning that the site is present. This digestion can be found in Figure 5-11.

\[\text{Figure 5-11. pPVU1t2 verification via plasmid digestion. Lane 1 is 200 ng of undigested plasmid (green arrows); lane 2 is 200 ng of pPVU1t2 digested with BamHI (5486 bp; blue arrow); lane 3 is 200 ng of pPVU1t2 digested with PvuII (5486 bp; blue arrow). The ladder (M) is GeneRuler DNA Ladder Mix (NEB).}\]

\[
\text{pPVU1t2 was transformed in } E. \text{ coli HB101 and the same methyltransferase activity test as with pACYC184-derived plasmids (Figure 5-8) was performed. The resulting FIGE gel can be observed in Figure 5-12.}
\]

\[\text{Figure 5-12. FIGE to test methyltransferase activity test in vivo for pPVU1t2. M1: λ DNA·HindIII (NEB); M2: λ DNA Monocut Mix (NEB); 1: E. coli HB101 genomic DNA plug; 2: HB101 genomic plug, PvuII-digested; 3: HB101 pPVU1t plug; 4: HB101 pPVU1t plug, PvuII-digested.}\]

In this case, methyltransferase action is confirmed: the pattern of the PvuII-digested control strain (lane 2) is not seen in the pPVU1t2 strain (lane 4), even though it is probably nicked or partly cleaved since it was mobile enough to migrate from the well.
Now that a strain with a working *Pvu* II methyltransferase has been obtained, the next step would be cloning in a different plasmid with *Pvu* II. pPVU1 has an incompatibility locus homologous to that of CoIE1 (Calvin Koons and Blumenthal, 1995), meaning that plasmids with pUC, pMB1 and pBR322 origins of replication are not compatible with it. pQR445, which contains a RSF1010 ori, should therefore be able to coexist in a cell with pPVU1t, so it was again decided to attempt to clone *Pvu* II into pQR445, but this time transforming it into competent HB101 pPVU1t cells. Nevertheless, attempts to even transform the naked pQR445 vector were unsuccessful and no colonies were isolated, suggesting that both plasmids are incompatible. This approach was ultimately abandoned in favour of a dual recombinase expression system.

5.2.2 Using a dual recombinase expression system

Recombinase expression facilitates better control of expression than simple inducible promoters. Lysogenic bacteriophages integrate their genome within the host genome as part of their infective cycle so that it propagates across many generations, until it enters a lytic state and the prophage excises from the host DNA. Phage-encoded integrases and excisionases mediate these steps: integrases are recombinases that recognise attB (host) and attP (phage) sequences and promote recombination between them, resulting in sites attL and attR. These site-specific recombinases are grouped in two families: serine or tyrosine recombinases, depending on the nucleophilic residue in the active site.

The dual plasmid system used in this work is derived from a series of plasmids developed by the Endy lab using *E. coli* DH5αZ1 (Bonnet et al., 2012, 2013). In these plasmids, the *araB* and PLTetO1 promoters control expression of Bxb1 and TP901 integrases through tetracycline and L-arabinose induction respectively as depicted in Figure 5-13. TP901 was selected in order to be able to utilise this system in a wider variety of hosts, since the PLTetO1 promoter is leaky in a tetR strain whereas the only requirement for a strain to utilise the *araB* promoter is to have an arabinose transporter. These plasmid maps can be found in the Appendix, sections 8.1.16 and 8.1.17.
5.2.2.1 Testing the sensitivity of the dual recombinase expression system

In order to evaluate the sensitivity and leakiness of this system in the absence of inducer, *E. coli* DH5αZ1 harbouring the dual recombinase system was grown and subjected to different L-arabinose concentrations in triplicate in a 96 well plate. mkate2 expression as a result of L-arabinose induction and OD\textsubscript{600} were monitored in a plate reader (see Chapter 2, section 2.4.7 for more details). The results are displayed in Figure 5-14.

**Figure 5-14.** Graph depicting relative fluorescence units (RFU) normalised by OD\textsubscript{600} in *E. coli* DH5αZ1 harbouring the dual recombinase system and induced at different L-arabinose concentrations to trigger mkate2 expression. 200 µL of cell culture at the beginning of exponential phase was transferred in triplicates to 96 flat-well plates containing the appropriate inducer volumes and transferred to a TECAN Infinite 200 PRO plate reader at 37 °C. Absorbance at 600 nm was measured every 15 min over 13 h and plates were shaken before measurement.

No expression is detected in the absence of inducer, meaning that this system is tightly repressed in absence of L-arabinose. Fluorescence is clearly observed at 1 % and 0.1 %
L-arabinose, where fluorescence starts to be observed 3 – 4 h after induction, while L-arabinose at 0.01 % yields a small amount of expression. This indicates that this recombinase-controlled system may be suitable for lethal protein expression such as restriction enzymes without the need for a strain expressing the cognate methyltransferase.

5.2.2.2 Cloning DSB-inducing proteins using the dual recombinase expression system

Given the absence of suitable restriction sites for cloning in the expression plasmid, the vector and the *Pvu* II CDS were PCR-amplified to introduce *Bam*HI and *Xho*I sites and enable cloning in place of the mkate2 CDS. Part of the mkate2 sequence was included to avoid PCR amplifying directly from the att regions, which contain highly repetitive sequences. A diagram illustrating the cloning strategy can be found in Figure 5-15.
Figure 5-15. Scheme depicting the cloning strategy for DSB-inducing proteins into the controller plasmid of the dual recombinase system. The backbone is obtained through PCR amplification of the vector minus mkate2 and introducing XhoI and BamHI sites instead. The PvuII sequence is amplified from pPVU1 introducing the same XhoI and BamHI sites. Both are digested and ligated, assembling pQR1872, in which PvuII is flanked by attB and attP sites to be recognised by L-arabinose-inducible TP901 integrase coded in the controller plasmid (not shown in image; see Figure 5-13).

The agarose gel for backbone amplification can be seen in Figure 5-16, and the gel for PvuII amplification can be seen in Figure 5-17. More details on the cloning procedure can be found in Chapter 2, section 2.2.6.8.
Figure 5-16. PCR of the expression vector backbone introducing BamHI and XhoI sites (5938 bp; lane 1, blue arrow). The orange arrow indicates a non-specific amplicon at 400 bp. The backbone was gel-purified. 10 µL of the PCR reaction were loaded. The ladder (M) is GeneRuler DNA Mix (NEB).

Figure 5-17. PCR of the PvuII CDS (506 bp; lane 1, blue arrow). The orange arrow indicates a non-specific amplicon at 400 bp. The fragment was gel-purified. 5 µL of the PCR reaction were loaded. The ladder (M) is GeneRuler DNA Mix (NEB).

PvuII was the first DSB-inducing protein to be successfully cloned in the expression plasmid of the dual system, yielding plasmid pQR1872. In this instance, screening was done via miniprep digestion, the gel for which is displayed in Figure 5-18.
Figure 5-18. Screening of expression plasmids containing the PvuII CDS via miniprep digestion with BamHI and XhoI. 1 µg of digested plasmid were loaded in each lane. Negative plasmids are uncut (lanes 1, 3, 7; orange squares) while positives yield 5912 and 480 bp fragments (lanes 2, 6, 8; blue and green arrows respectively). Lane 4 is unclear, since it does not show a 480 bp fragment but has been clearly linearised. The ladder (M) is the GeneRuler DNA Ladder Mix (Thermo Fisher).

Attempts to co-transform of both controller plasmid and pQR1872 were unsuccessful, and no colonies were observed. Instead, a sequential transformation protocol was required: competent cells of DH5αZ1 harbouring the controller plasmid had to be prepared, and then transformed with pQR1872. These cells were then grown and PvuII activity tested, as seen in Figure 5-19, where incubation of λ DNA with a clarified lysate of an induced cell sample results in a similar pattern to digestion with commercial PvuII. The induced cells grew to a lower OD$_{600}$ than the non-induced, which will be discussed later.

Figure 5-19. In vitro activity test of PvuII in E. coli DH5αZ1 harbouring the controller plasmid and pQR1872. Lane 1: λ DNA digested with commercial PvuII; lane 2: λ DNA with clarified lysate from a non-induced sample; lane 3: λ DNA with clarified lysate from an induced sample.
Once this dual integrase system was confirmed to be suitable for *Pvu*II expression, five other proteins were selected to be tested in the genome ablation cassette: *Not*I, *Bam*HI, *Hpa*II, denA, and ccdB. This time, the restriction enzymes were cloned without their cognate methyltransferases. The restriction enzyme genes were PCR-amplified from genomic DNA samples, while denA and ccdB were ordered as gBlocks (the DNA sequences can be found in the Appendix, sections 8.4.1 and 8.4.2). Screening was done via colony PCR given the large number of proteins to clone; an example gel is pictured in Figure 5-20.

![Figure 5-20. Colony PCR screening verifying CDS insertion in the expression plasmid. The oligonucleotides targeted the attB and attP sites. Lanes 1-5 are for *Pvu*II; lanes 6-10 for *Not*I; lanes 11-12 for *Bam*HI; lanes 13-17 for *Hpa*II; lanes 18-22 for ccdB; and lanes 23-27 for endonuclease denA. The different bands are discussed in-text. The whole colony PCR reactions (10 µL) were loaded. The ladder (M) is the 100 bp DNA ladder (NEB).](image_url)

The colony PCR results were unexpected, since the amplicons were a different size than expected, or varying sizes. *Pvu*II was included to verify that colony PCR yielded the expected insert (737 bp, blue arrows in lanes 1-5), and is one of the few cases in which it did. While a 1415 bp insert was expected for *Not*I, the amplicons ranged from 200 to 400 bp (lanes 6-10, orange squares); for *Bam*HI, a 905 bp fragment was expected, and instead 400 and 600 bp fragments were obtained (lanes 11-12, orange squares); for *Hpa*II, 400 and 700 bp fragments were obtained instead of the expected 1340 bp (lanes 13-17, orange squares); for ccdB, one of the fragments matched the expected 569 bp (lane 21, blue arrow), while the rest were slightly lower at 500 bp (lanes 18, 19, 20, 22;
orange squares); and 350 and 1500 bp fragments were obtained for denA instead of the expected 698 bp (lanes 23-27, orange squares). None of the unexpected fragments match the size that would be observed with a naked plasmid as template for the colony PCR, 827 bp. A gel was run to confirm that all inserts were indeed the expected size post-digestion, which can be found in Figure 5-21.

![Figure 5-21. Gel depicting 50 ng of the different inserts ready to clone into the expression plasmid, indicated with a blue arrow: PvuII (lane 1, 486 bp), NotI (lane 2, 1164 bp), BamHI (lane 3, 654 bp), HpaII (lane 4, 1091 bp), denA (lane 5, 318 bp) and ccdB (lane 6, 449 bp). The ladder (M) is the 100 bp DNA ladder (NEB).](image)

In some cases, miniprep digestion can be a more reliable screening method than colony PCR. Therefore, one colony for each candidate was picked, and minpreps prepared and digested with XhoI and BamHI like in Figure 5-18. This gel is depicted in Figure 5-22.

![Figure 5-22. Miniprep digestions of the colonies belonging to lanes 10, 11, 15, 21 and 25 screened in Figure 5-20, in order (NotI, BamHI, HpaII, ccdB and denA respectively). 1 µg of digested plasmid was loaded. The green squares indicate the backbone. The different bands highlighted with blue arrows are discussed in-text. The ladder used is the 100 bp DNA ladder (NEB).](image)

The right insert was only obtained for ccdB (306 bp; blue arrow, lane 4). The rest yielded inserts of the wrong size: no insert was obtained when digesting the plasmid DNA of the NotI colony (lane 1), where a 1158 bp band would be expected; 1517 bp and 400 bp
inserts were obtained for *Bam*HI (blue arrows, lane 2) instead of the expected 642 bp; a 400 bp insert was obtained for *Hpa*II (blue arrow, lane 2), while a 1077 bp was expected; and a 100 bp and 1000 – 1200 bp insert were obtained in denA, while a 435 bp band was expected (blue arrows, lane 5). These inserts do not match the amplicon sizes obtained through colony PCR either, except for ccdB, where the one positive colony had the right amplicon (lane 21 in Figure 5-20).

Since the recombinase expression system was the only method that enabled *Pvu*II expression and cloning was successful in two instances, more colonies were screened until the right fragment was obtained. However, the cloning difficulties did not end here. When more colonies were screened via colony PCR, the first positive one for *Hpa*II with the right size amplicon turned out to not be active when testing it for *in vitro* activity and no OD$_{600}$ decrease was observed like in the case of *Pvu*II. More colonies with the right size were screened until one showed both reduced growth post-induction and the activity could be seen in a gel, which is observed in Figure 5-23.

![Figure 5-23. In vitro activity test of HpaII in E. coli DH5αZ1 harbouring the controller plasmid and pQR1873. Lane 1: λ DNA digested with commercial HpaII; lane 2: λ DNA with clarified lysate from a non-induced sample; lane 3: λ DNA with clarified lysate from an induced sample.](image)

Activity tests for the restriction enzymes were straightforward, since they simply involved restriction *in vitro* in the right buffer and visualising the result in an agarose gel. However, in the case of ccdB this proved more troublesome, since it is a toxin that interferes with
DNA gyrase activity. Moreover, despite different buffers were tested, *in vitro* denA assays did not show digestion. Therefore, the screening strategy changed: after obtaining the correct fragment via colony PCR, plasmid DNA was isolated and transformed into DH5αZ1 with the controller plasmid. The resulting strain was grown to exponential phase, a non-induced and an induced culture in parallel, and OD<sub>600</sub> was recorded 4 h post-induction. If OD<sub>600</sub> was observed to be significantly lower in the induced culture, the cloning was deemed successful. Ultimately, four plasmids were obtained: pQR1872 (PvuII), pQR1873 (HpaII), pQR1874 (ccdB) and pQR1875 (denA). Their plasmid maps can be found in the Appendix, sections 8.1.19 to 8.1.22. The registered OD<sub>600</sub> post-induction can be observed in Table 5-1.

Table 5-1. OD<sub>600</sub> measurements of different DH5αZ1 cultures harbouring the controller plasmid and the different endonuclease plasmids 4 h after induction. Both cultures were inoculated with the same amount of starter and were at an OD<sub>600</sub> = 0.6 – 0.8 at the time of induction.

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Non-induced</th>
<th>Induced</th>
</tr>
</thead>
<tbody>
<tr>
<td>pQR1872 (PvuII)</td>
<td>2.45</td>
<td>0.935</td>
</tr>
<tr>
<td>pQR1873 (HpaII)</td>
<td>2.745</td>
<td>1.975</td>
</tr>
<tr>
<td>pQR1874 (ccdB)</td>
<td>1.025</td>
<td>0.49</td>
</tr>
<tr>
<td>pQR1875 (denA)</td>
<td>2.61</td>
<td>1.27</td>
</tr>
</tbody>
</table>

After ascertaining that these plasmids indeed encoded working proteins, they were sent for sequencing using primers #S-03 and #S-04 (sequence available in the Appendix, section 8.2.1). More unexpected results were obtained: one frame-shifting deletion was found in the PvuII and ccdB sequences, despite protein activity was confirmed via just a decrease in OD<sub>600</sub> 4 h after induction (ccdB and denA) or via OD<sub>600</sub> decrease and *in vitro* activity tests (PvuII). pQR1872 was sent again with an additional two sequencing primers for each strand (oligonucleotides #S-05 to #S-08; sequence available in the Appendix, section 8.2.1) but this mutation was present in all readings and introduced early stop codons as depicted in Figure 5-24.
Alternative start codons that would maintain the correct reading frame were not found upstream of the mutation. One alternative initiation codon, TTG, was found downstream of the mutation, maintaining the original reading frame and resulting in an ORF that is missing 38 codons at 5’ (PvuII is 157 residues long). This alternative ORF is depicted in Figure 5-25.
Therefore, it seems possible that this alternative codon is being recognized and leads to translation of an incomplete but functional \textit{PvuII}. It is also interesting that this type of mutation was found for \textit{PvuII} but no mutations were found in pQR1875, encoding \textit{denA}, or pQR1873, encoding \textit{HpaII}. Both are higher frequency cutting endonucleases than \textit{PvuII}, with an estimated 57,096 and 24,311 cleavage sites respectively in the \textit{E. coli} MG1655 genome while \textit{PvuII} has an estimated 1,776 sites.

In the case of the \textit{ccdB} CDS, in this case the frame-shifting deletion resulted in \textit{ccdB} with 20 different amino acids at C-terminal, and a protein 6 amino acids smaller overall. The sequence alignment highlighting the deletion in \textit{ccdB} is displayed in Figure 5-26.
Figure 5-26. Snapshot of SnapGene showing a portion of the alignment of the original designed sequence of the beginning of the ccdB CDS in pQR1875 (above, purple) and the sequencing results, depicting the same frame-shifting deletion near the end of the sequence (red square), and the resulting protein sequence with a premature stop codon (grey).

It is unclear why cloning *Pvu*II was straightforward, but the other proteins were more troublesome: the cloning was performed in a DH5αZ1 strain lacking the recombinase expression plasmid, and without it, the genes in the expression plasmid would be cloned in the reverse orientation. In other words, no expression whatsoever should have been observed, but the colony screening results are a typical consequence of the selective pressure exercised through an attempt at cloning toxic proteins in a system that does not suppress expression effectively under non-induced conditions. Attempts at cloning restriction enzymes have been reported in which unexpected toxicity was observed (Bao et al., 2008), which may be attributed to the presence of a cryptic promoter or leaky expression from the tac promoter. Other instances of unexpected expression have been observed during the design of genetic circuits. In one case antisense transcription was observed due to a *araB* promoter with antisense transcription sites, which was solved via the introduction of a stronger terminator (Gorochowski et al., 2017). It this case, it could be possible that some sites in the plasmid have undescribed transcription initiation sites or promoters that might be activated under certain conditions and allow for transcription in such low quantities that cause a minimum amount of toxicity sufficient to act as selective pressure. In addition, it has been reported that not only AUG, GUG and UUG are capable of translation initiation, but other 44 codons are also capable of doing
so, albeit in lower proportions, which could also be exerting this effect (Hecht et al., 2017).

5.2.2.3 Cloning the exonuclease cassette into the dual recombinase expression system

DSB-inducing protein expression requires the presence of two plasmids in one cell: one Cam-resistant (the controller plasmid) and another Amp-resistant (the expression plasmid). The exonuclease cassette, composed of exonuclease III and T5 exonuclease cloned under the tac promoter in pTTQ18 as explained in Chapter 3, had to be cloned into either of these vectors for two reasons: pTTQ18 is also Amp-resistant, like the controller plasmid, and having two plasmids resistant to the same antibiotic in one cell does not guarantee maintenance of both; and having three different plasmids in one cell would a significant burden. Therefore, the whole cassette expressing lacI together with its promoter, the tac promoter, both exonuclease genes and terminator were PCR-amplified to introduce XhoI and BamHI sites and enable ligation into the controller plasmid, resulting in a plasmid termed pQR1871. Details of the cloning procedure can be found in Chapter 2, section 2.2.6.7. The gel depicting the PCR can be observed in Figure 5-27.

After transformation of the ligation, colonies were picked, and screening was done via miniprep digestion with XhoI and SmaI. The gel for the screening is pictured in Figure 5-27.

Figure 5-27. PCR amplification of the exonuclease cassette (4327 bp, blue) prior to cloning into the controller plasmid. Due to the presence of non-specific bands (orange), the fragment was gel purified. 5 µL of the PCR reaction were loaded. The ladder (M) is the GeneRuler DNA Ladder Mix (NEB).
One positive colony was selected, sent for sequencing to confirm the right sequences for the exonuclease, and the plasmid was termed pQR1871. The plasmid map can be found in the Appendix, section 8.1.18.

Figure 5-28. Colony screening of pQR1871. 1 µg of digested plasmid were loaded. Positive colonies (lanes 1, 4, 5) yield a 5524 bp and 4303 bp band (blue and green arrows respectively), whereas negative colonies yield 5524 bp (orange square), 783 bp and 306 bp bands. The latter two are not observed on the gel because it was run for an excessive length of time and they must have run out of the gel. Lane 8 appears to be undigested plasmid (purple square).

5.2.3 Ascertaining the optimal induction time

In molecular biology experiments, cultures are routinely induced at OD₆₀₀ = 0.6 – 0.8, when cells are starting to grow exponentially and are healthy and very metabolically active. However, this is not necessarily the optimal point of induction. In batch fermentation, induction is routinely performed when the dissolved oxygen level increases suddenly (or ‘spikes’), which signals that the carbon source has been completely consumed cells are no longer growing exponentially, therefore not requiring as much oxygen (Wehrs et al., 2020). Larentis et al. (2014) found higher productivity for immunoglobulin-like protein in shake flask growth when inducing at OD₆₀₀ = 2. When expressing toxic proteins, inducing later in the growth phase may be preferable: the use of stationary phase promoters, which are activated when cell density is high, is an alternative in these cases, since an inducer is not required and high production of toxin will not be compromising cell growth. Cao and Xian (2011) compared the yield of phloroglucinol, a bulk chemical commonly used in the pharmaceutical industry, by cloning gene phlD responsible for phloroglucinol biosynthesis in two expression systems:
the IPTG-inducible T7 promoter, and the stationary phase promoter of the *fic* gene, a cell division regulatory factor. Despite being weaker than T7, phloroglucinol yield was higher when using the *fic* promoter.

In these examples, the goal is to maximise the yield of a product by expressing as much protein as possible, whereas here the product are the cells themselves and it cannot be assumed that later induction will be more efficient. For this reason, it was deemed necessary to ascertain whether induction at the beginning or the end of the exponential phase was optimal in terms of DNA content and viable cell count. *E. coli* DH5αZ1 pQR1871 pQR1872 was grown in shake flasks to early exponential phase; then, different conditions were tested in duplicate: non-induction; IPTG induction only, at either 3 or 7 h; L-ara induction only, at either 3 or 7 h; sequential IPTG and L-ara induction at 3 h and 7 h respectively; sequential L-ara and IPTG induction at 3 h and 7 h respectively; and simultaneous IPTG and L-ara induction, at either 3 or 7 h. 1 mL samples were taken for DNA extraction (for more details see Chapter 2, section 2.2.1.11) and 250 µL samples for cfu/mL determination to calculate the escape frequency (for more details see Chapter 2, section 2.4.6). This experiment was performed with pQR1872 because at this point of the project *PvuII* was the only DSB-inducing protein successfully expressed, and this work was done in parallel to the rest of cloning described in section 5.2.2.2. The growth curves obtained for this experiment can be found in Figure 5-29.
Figure 5-29. Growth curves of DH5αZ1 pQR1871 pQR1872 at different induction conditions. Each condition was run in duplicate in 250 mL flasks grown at 37 °C and 250 rpm. The time in parenthesis is the time after inoculation at which that element was induced; for example, ‘PvuII (3 h)’ means that PvuII expression was induced via addition of L-ara at 3 h post-inoculation; ‘PvuII (3 h), exo (7 h)’ means that PvuII expression was induced at 3 h, and exonuclease expression was induced at 7 h; ‘PvuII, exo (3 h)’ means that both PvuII and exonuclease expression were induced at 3 h post-inoculation. ‘n/Ind’ is the control with no induction. The error bars indicate the range between the minimum and maximum absorbance values. The time of induction is not indicated with an arrow due to the different induction times for different samples.

As expected, the non-induced sample has the highest OD$_{600}$ throughout the experiment, followed by those samples in which the first induction occurred at 7 h, with PvuII induction, and last, those in which the first induction occurred at 3 h. In this last group, the decrease in OD$_{600}$ is not as substantial when only inducing exonuclease expression at 3 h (both exonucleases only, and followed by PvuII induction at 7 h), PvuII induction causing a larger reduction. Moreover, the lowest OD$_{600}$ is obtained when inducing both PvuII and exonucleases at 3 h.
Escape frequencies were also calculated in this experiment. They are a measure of how easily cells overcome lethality safeguards, such as kill switches, and are an important measure of biocontainment efficiency: the National Institutes of Health (NIH) in the US state in their guidelines for research involving recombinant or synthetic nucleic acids that a desirable escape frequency is $10^{-8}$ via survival of the host organism or transmission of the nucleic molecule to other organisms (National Institutes of Health, 2016). This value is often cited and compared to when testing kill switches and other genetic safeguards (Moe-Behrens et al., 2013; Torres et al., 2016). Escape frequencies are calculated by dividing the cfu/mL grown at non-permissive conditions (in this case, upon nuclease induction) divided by the same value in permissive conditions (a non-induced control). The escape frequencies obtained in this experiment are displayed in Figure 5-30.
Figure 5-30. The graph shows the escape frequencies measured at different times after inoculation for the different induction conditions. The time in parenthesis is the time after inoculation at which that element was induced; for example, 'PvuII (3 h)' means that PvuII expression was induced via addition of L-ara at 3 h post-inoculation; 'PvuII (3 h), exo (7 h)' means that PvuII expression was induced at 3 h, and exonuclease expression was induced at 7 h; 'PvuII, exo (3 h)' means that both PvuII and exonuclease expression were induced at 3 h post-inoculation. Escape frequencies are calculated as the cfu / mL obtained for each induced sample divided by the same value for the control, non-induced cultures. The error bars indicate the range between the minimum and maximum absorbance values.
The escape frequency at 3 h is not included in the figure, since non-permissive conditions are just introduced at that timepoint and it was not deemed worthwhile to calculate them for all samples because they should be 1. They were calculated for two samples (the sequential induction cultures) and found to be 0.95 ± 0.12 and 1 ± 0.05. This is also confirmed in those samples in which induction happens at 7 h (Pvull 7 h, exo 7 h, and Pvull, exo 7 h), where the escape frequencies are approximately 1.

Unfortunately, the escape frequencies obtained are far from the desired $10^8$, the lowest one being $10^4$. The lowest frequencies are obtained when inducing Pvull at 3 h (for both Pvull only, and sequential induction with Pvull at 3 h) up to 11 h after inoculation: at 24 h the escape frequency has increased, meaning that cells are overcoming the deleterious effect of nuclease expression. The only instance in which this increase does not happen is when the first induction occurs at 7 h (Pvull only, exonucleases only, or both), for which escape frequencies actually decrease slightly.

Induction of both elements simultaneously is more efficient at 7 h than at 3 h: in addition to the escape frequency decreasing after 24 h, at 11 h their escape frequency is lower than for cells induced at 3 h, especially remarkable given that for these cells an additional 4 hours have elapsed since induction.

Interestingly, when inducing one element only at 3 h, Pvull induction causes a lower escape frequency than exonuclease induction at 7, 11 and 24 h; this reverts when inducing at 7 h, where exonuclease induction causes a lower escape frequency than Pvull. In addition, when looking at sequentially induced samples, first inducing Pvull and then the exonucleases seems a better option than first expressing exonucleases and then the endonuclease, even though eventually both sequential induction samples reach the same escape frequency after 24 h. This suggests that expression of a restriction enzyme in the early exponential phase is more deleterious to growth than exonucleases, whereas these may be more harmful than the endonuclease when induced later.
Nuclease expression alone is not enough to achieve an acceptable escape frequency of $10^{-8}$. However, as seen in Chapter 4, the number of viable cells can be reduced to zero when adding MMC, so this issue can be solved easily by adding a DSB-inducing antibiotic. Therefore, the decision on which induction conditions are optimal should be done based on DNA content reduction instead.

The DNA content estimation was performed via densitometry in the same manner as in Chapter 4, sections 4.2.2 and 4.2.3. The DNA gels can be found in the Appendix, section 8.6, Table 8-8. The variation in DNA content is calculated as the DNA amount at a timepoint divided by the DNA content at time of induction, and in this experiment, induction occurs at either 3 h or 7 h; therefore, the results for both subsets are displayed in separate graphs and can be found in Figure 5-31. In addition, given the large discrepancy between the non-induced control and the rest of samples, the subset induced at 3 h is portrayed in Figure 5-32 to better visualise the results.
Figure 5-31. Variation in DNA content of samples with different induction conditions in which the first induction occurs at 3 h (left) or at 7 h (right). Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h or 7 h, in the same agarose gel. The time in parenthesis is the time after inoculation at which that element was induced; for example, 'PvuII (3 h)' means that PvuII expression was induced via addition of L-ara at 3 h post-inoculation; ‘PvuII (3 h), exo (7 h)’ means that PvuII expression was induced at 3 h, and exonuclease expression was induced at 7 h; ‘PvuII, exo (3 h)’ means that both PvuII and exonuclease expression were induced at 3 h post-inoculation. ‘n/Ind’ is the control with no induction. The error bars indicate the range between the minimum and maximum values.

The DNA gels used for these calculations can be found in the Appendix, section 8.6, Table 8-8.
Results from Figure 5-31 (left) without the non-induced control. Results are expressed as the intensity of the DNA band pertaining to each timepoint normalised by the DNA content at the time of induction, 3 h, in the same agarose gel. The time in parenthesis is the time after inoculation at which that element was induced; for example, ‘PvuII (3 h)’ means that PvuII expression was induced via addition of L-ara at 3 h post-inoculation; ‘PvuII (3 h), exo (7 h)’ means that PvuII expression was induced at 3 h, and exonuclease expression was induced at 7 h; ‘exo (3 h), PvuII (7 h)’ means that both PvuII and exonuclease expression were induced at 3 h post-inoculation. The error bars indicate the range between the minimum and maximum values. The DNA gels used for these calculations can be found in the Appendix, section 8.6, Table 8-8.

In the first subset of samples with a first induction at t = 3 h, the non-induced control displays a large increase in DNA content. The rest of conditions achieve a similar amount of DNA content, which in the case of simultaneous induction of PvuII and exonucleases at 3 h increases over time. Exonuclease only induction at 3 h does not seem to cause a great DNA content reduction by t = 7 h, which at first instance would seem to contradict the results obtained in Chapter 3. However, those results cannot be extrapolated since in this case the exonuclease cassette is in a different vector and a different strain alongside another plasmid.
In the second subset with a first induction at $t = 7$ h, there is not a great difference in DNA content with respect to the non-induced sample, suggesting that induction at 3 h is a better option. In this case, simultaneous induction of $Pvul$ and exonucleases at 7 h is more effective than $Pvul$ or exonuclease induction only. It is also interesting to observe that $Pvul$ induction at 7 h yields a similar DNA content than the non-induced control, despite $Pvul$ induction at 7 h causing a reduction in both $OD_{600}$ and escape frequency as seen in Figure 5-29 and Figure 5-30 respectively.

The ideal conditions in terms of DNA reduction would be induction of $Pvul$ and exonucleases at 3 h or sequential induction of $Pvul$ (3 h) and exonucleases (7 h), the latter staying consistently just below 40% throughout the experiment. There does not seem to be much difference between both conditions at the 7 h timepoint, 4 h after the first induction; therefore, simultaneous induction at 7 h was selected as the most optimal condition due to an easier experimental procedure.

### 5.2.4 Testing different DSB-inducing proteins

Once the optimal conditions for DNA reduction were concluded to be simultaneous induction of DSB-inducing protein and exonuclease early in the exponential phase, the four different proteins successfully cloned and expressed are tested in *E. coli* DH5αZ1 together with exonuclease III and T5 exonuclease expression. To obtain these strains, like in section 5.2.2.2, first DH5αZ1 pQR1871 competent cells were prepared and then the different plasmids were transformed, since simultaneous transformation of both plasmids yielded no colonies. *E. coli* DH5αZ1 with plasmids pQR1871 and pQR1872 (*Pvul*), pQR1873 (*HpaII*), pQR1874 (ccdB) or pQR1875 (denA) was grown in triplicates in 250 mL shake flasks, induced at 3 h and grown for a total of 24 h. The control in this case was this strain harbouring the naked recombinase controller and expression plasmids. The growth curves can be observed in Figure 5-33.
Figure 5-33. Growth curves of E. coli DH5αZ1 expressing exonucleases (pQR1871) together with a DSB-inducing protein: PvuII (pQR1872), HpaII (pQR1873), ccdB (pQR1874), or denA (pQR1875). Cultures were grown in triplicates in 250 mL shake flasks, induced at 3 h and grown for 24 h at 37 °C and 250 rpm. The blue arrow indicates the point of induction, at 3 h post-inoculation. The error bars indicate the standard deviation.

As expected and observed previously, OD$_{600}$ stops increasing upon induction, indicating that nucleases or other DSB-inducing proteins are being expressed. This is especially noticeable in the case of HpaII (pQR1873).

Escape frequencies are not calculated in this case due to the different conditions in this experiment: in section 5.2.3, the same strain, DH5αZ1 harbouring the naked controller and expression plasmids, was grown under different conditions, whereas here different strains are being compared (DH5αZ1 pQR1871 with either pQR1872, pQR1873, pQR1874 or pQR1875). Therefore, to evaluate reduction of viability, the metric selected is the ratio of the cfu / mL obtained at 7 h and 24 h divided by cfu / mL at 3 h. This is displayed in Figure 5-34.
Figure 5-34. Ratio of the viable cell count obtained at 7 h and 24 h after inoculation divided by the viable count at the time of induction, 3 h after inoculation, for different DH5αZ1 expressing exonucleases (pQR1871) and DSB-inducing proteins PvuII (pQR1872), HpalII (pQR1873), ccdB (pQR1874), or denA (pQR1875) alongside a non-induced control. The viable cell count was obtained by sampling triplicate cultures at 37 °C and 250 rpm over 24 h and plating on LB agar. The error bars indicate the standard deviation.

The lowest ratio is achieved at 7 h, 4 h after HpalII induction (pQR1873), followed by ccdB (pQR1874). There is an increase in viability, herein expressed via an increase in cfu / mL, for all DSB-inducing proteins after 24 h, which mirrors what was observed when inducing both PvuII and exonucleases simultaneously at 3 h in DH5αZ1 pQR1871 pQR1872 (see Figure 5-30).

There is not a great difference in terms of reduction of cfu, although as mentioned in section 5.2.3, viability ceases to be an issue if an additional DSB-inducing antibiotic is added to the culture. Therefore, in this case, DNA content is also deemed more crucial than reduction of cfu / mL. Like in the previous experiment, the reduction in DNA content is measured as the total DNA at each timepoint, taking the DNA content at the time of induction as 100 %; the only difference was that, in this case, DNA was quantified via
the PicoGreen assay. The results can be found in Figure 5-35. Figure 5-36 is the same figure minus the control, to enable better visualisation and interpretation of results.

![DNA content relative to pre-induction values](image)

**Figure 5-35. Analysis of total DNA content in E. coli DH5αZ1 expressing exonucleases (pQR1871) and DSB-inducing proteins PvuII (pQR1872), HpaII (pQR1873), ccdB (pQR1874), or denA (pQR1875) alongside a non-induced control. Results are expressed as the DNA amount as measured by the PicoGreen assay at each timepoint normalised by the DNA content at the time of induction, 3 h. The error bars indicate the standard deviation. The DNA gels can be found in the Appendix, section 8.6, Table 8-9.**
Figure 5-36. Analysis of total DNA content in E. coli DH5αZ1 expressing exonucleases (pQR1871) and DSB-inducing proteins PvulI (pQR1872), HpalI (pQR1873), ccdB (pQR1874), or denA (pQR1875) minus the non-induced control. Results are expressed as the DNA amount as measured by the PicoGreen assay at each timepoint normalised by the DNA content at the time of induction, 3 h. The error bars indicate the standard deviation. The DNA gels can be found in the Appendix, section 8.6, Table 8-9.

As expected, the control displays an expected large increase in DNA content at both 7 h and 24 h. The effect of nuclease expression in DNA content is pronounced, especially in the case of HpalI (pQR1873) at 7 h, even though after 24 h DNA content increases to similar values to the other strains. This may have to do with the characteristics of these DSB-inducing proteins: PvulI and HpalI are restriction enzymes, with an estimated 1,776 and 24,311 cleaving sites in the E. coli MG1655 genome. A small amount of HpalI would be able to cleave a genome in many more sites than the equivalent amount of PvulI, which explains the larger DNA reduction and lower viability 4 h after induction (t = 7 h). However, it is surprising that denA is not more efficient, since it is an endonuclease with a CCGC sequence specificity and an estimated 57,096 sites in the E. coli MG1655 genome (Carlson et al., 1999). As for ccdB, it is a DNA gyrase poison that traps the
enzyme-DNA cleavage complex, so it is susceptible to cause as many DSBs as replication and transcription sites are in a chromosome, a number difficult to estimate given that it depends on the metabolic state of the cell.

HpaII induction at 3 h is optimal for DNA reduction if cells are sampled at 7 h. These experiments were performed on E. coli DH5αZ1, since this is the model strain in which the recombinase-mediated expression system was developed. However, E. coli CSR603 could potentially be a better candidate: this strain has the genotype recA1 uvrA6 phr-1 (mutations in recA, for DSB repair; uvrA, for nucleotide excision repair; and phrB, for photoreactivation repair). Its consequent inability to repair DNA damage made it the strain of choice for generation of Maxicells (Sancar et al., 1979). In contrast, in terms of mutations in DNA repair genes, DH5αZ1 only contains the recA1 mutation. Therefore, even though genome ablation cannot be accomplished in DH5αZ1 through protein expression, there was a possibility that it could occur in CSR603.

5.2.5 Comparing DH5αZ1 and CSR603

The original idea for this experiment was to express the four different DSB-inducing proteins in E. coli CSR603 and compare them to E. coli DH5αZ1. However, only PvuII was successfully expressed in CSR603: despite colonies were obtained when transforming plasmids pQR1873, pQR1874 and pQR1875 in competent E. coli CSR603 pQR1871, when these were grown and induced, OD600 did not decrease and no activity was observed in in vitro assays of HpaII, even though different transformants were screened. Therefore, it was decided to compare the best performing combination in DH5αZ1 (with plasmids pQR1871 pQR1873, expressing HpaII) with CSR603 harbouring plasmids pQR1871 pQR1872 (expressing PvuII), alongside their respective controls with the plasmids with no nuclease genes.

Like in section 5.2.4, CSR603 pQR1871 pQR1872 was grown in triplicates to early exponential phase and induced, and DNA content and cfu / mL were recorded 4 h after
induction and after an overnight incubation. CSR603 grows more slowly than DH5αZ1, so in this case induction happens at 5 h. The growth curves are displayed in Figure 5-37. As expected, the induced nuclease-expressing cultures display a lower OD$_{600}$ than the controls, even though this is not as pronounced in CSR603.

![Graph showing growth curves of DH5αZ1 and CSR603 with and without QR1871 pQR1873 and QR1871 pQR1872.](image)

Figure 5-37. Growth curves of induced DH5αZ1 pQR1871 pQR1873 (HpaII, exonucleases) (left) and CSR603 pQR1871 pQR1872 (PvuII, exonucleases) (right) alongside their respective controls. In both cases, cultures were grown in triplicates in 250 mL shake flasks at 37 °C and 250 rpm for 24 h. The blue arrow indicates the point of induction. The error bars indicate the standard deviation.

Viability was also evaluated through the ratio between viable cell count obtained at 4 h post-induction and 24 h after inoculation divided by the viable count at the time of induction. This is displayed in Figure 5-38.
In both nuclease-expressing strains there is a decrease in viability as compared to their non-induced counterparts 4 h after induction, followed by a recovery up to similar levels to this control. A more drastic reduction in viability would be expected of CSR603 given its weakened DNA repair, but it is actually similar to DH5αZ1.

Like in the previous section, DNA content is considered more crucial. The decrease in DNA content is calculated as the total DNA at each timepoint, taking the DNA content at the time of induction as 100 %. Total DNA is also quantified via the PicoGreen assay. The results can be observed in Figure 5-39, and Figure 5-40 displays the induced samples only to enable better visualisation and interpretation of results.
Figure 5-39. Analysis of total DNA content induced E. coli DH5αZ1 pQR1871 pQR1873 (HpaII, exonucleases) and CSR603 pQR1871 pQR1872 (PvuII, exonucleases) alongside their respective controls with the empty plasmids. Results are expressed as the DNA amount as measured by the PicoGreen assay at each timepoint normalised by the DNA content at the time of induction, 3 h. The error bars indicate the standard deviation. The DNA gels can be found in the Appendix, section 8.6, Table 8-9.

Figure 5-40. Analysis of total DNA content induced E. coli DH5αZ1 pQR1871 pQR1873 (HpaII, exonucleases) and CSR603 pQR1871 pQR1872 (PvuII, exonucleases) minus their controls. Results are expressed as the DNA amount as measured by the PicoGreen assay at each timepoint normalised by the DNA content at the time of induction, 3 h. The error bars indicate the standard deviation. The DNA gels can be found in the Appendix, section 8.6, Table 8-9.
Surprisingly, not as much DNA was extracted from the CSR603 control as from the DH5αZ1 control. The non-induced strain was grown three times, and every time samples were taken and DNA quantified, but the same results were obtained. The only potential explanation is that during the extraction procedure native nucleases degraded the genomic DNA in CSR603 at a greater extent than for DH5αZ1 due to its genotype.

As would be expected, CSR603 is better performing than DH5αZ1, although only marginally: nuclease expression in CSR603 achieves a lower DNA content than the best DH5αZ1 combination (12.8 % and 22.4 % respectively). CSR603 pQR1871 pQR1872 is therefore considered a better candidate for genome ablation.

5.2.6 Addressing the increase in viability post-induction

Recovery of viability post-induction has been observed in most experiments so far. Ideally, a low level of nuclease production would be immediately lethal, but there are different methods in which cells overcome lethal switches. These include mutation of the lethal protein, plasmid loss, or outgrowth of plasmid-free cells naturally present in the population. A set of experiments were performed in order to ascertain the reason why cells overcome nuclease induction in this instance.

5.2.6.1 Testing restriction enzyme activity

The first experiment was to ascertain whether there was still detectable restriction enzyme activity after 24 h: an absence of activity would indicate that plasmid-free cells have outgrown the rest of the population, or that mutations have been developed. As mentioned in section 5.2.2.2, in vitro activity tests could only be performed for PvuII and HpaII. Therefore, only the strains expressing these restriction enzymes were used in this experiment: E. coli DH5αZ1 harbouring plasmids pQR1871 and either pQR1872 (PvuII) or pQR1873 (HpaII) and E. coli CSR603 harbouring pQR1871 and pQR1872 were grown until early exponential phase, induced and incubated for a total of 24 h. Samples were
taken at the end of this period and *in vitro* activity tests were performed as explained in Chapter 2, section 2.3.2.2, which can be observed in Figure 5-41.

![Image of gel showing restriction enzyme activity tests](image)

*Figure 5-41. In vitro restriction enzyme activity tests for E. coli DH5αZ1 pQR1871 pQR1872 (PvuII) (lane 1, non-induced; lane 2, induced), E. coli DH5αZ1 pQR1871 pQR1873 (HpaII) (lane 3, non-induced; lane 4, induced), and E. coli CSR603 pQR1871 pQR1872 (PvuII) (lane 5, non-induced; lane 6, induced). All extracts are incubated with λ DNA. Lane C1 is undigested λ DNA; lane C2 is λ DNA digested with commercial PvuII; lane C3 is λ DNA digested with commercial HpaII. If there was restriction activity present, lanes 2 and 6 would display a pattern similar to lane C2, the *Pvu*II-digested control; and lane 3 would display a pattern similar to lane C3, the *Hpa*II-digested control. However, this is not the case, meaning that there is no functional restriction enzyme present after 24 h in the culture.

There is a possibility that there are still cells present in the culture able to express these endonucleases, albeit in the minority. Therefore, a sample was reinoculated into fresh LB with antibiotics and this new culture was again grown until early exponential phase, induced and incubated for just 4 h. Samples were then taken and the *in vitro* activity tests were repeated, which can be found in Figure 5-42.
Only DH5αZ1 pQR1871 pQR1873 (HpaII) seems to retain some activity: all three induced replicates (lanes 4-6) display a similar pattern to the HpaII-digested control (lane C3), while the non-induced replicates do not display much λ DNA degradation. Both DH5αZ1 and CSR603 harbouring pQR1871 and pQR1872 (PvuII) do not display a different behaviour between non-induced (lanes 7-9 and 13-15) and induced (lanes 10-12 and 16-18) samples. This means that, at least in the case of DH5αZ1 pQR1871 pQR1873, some cells are still able to express HpaII despite they are not producing enzyme by the end of the first incubation period.
5.2.6.2 Verifying the presence of the controller and expression plasmids

Another possible explanation to cell recovery is the insertion of insertion elements, like transposon or insertion sequences, in either the controller or the expression vectors. Insertion sequences are short DNA sequences that can move to different positions in a cell genome and are suggested to have an evolutionary role, since they are a source of genetic variability.

In the next experiment, all strains encoding exonucleases and DSB-inducing proteins were grown in 250 mL flasks, induced, and the plasmid DNA was isolated after a total 24 h of growth. This was done in triplicate, alongside one non-induced control. 1 µg of plasmid DNA was digested with XhoI and BamHI, combination that would yield a pattern that would enable detection of both plasmids, and run on an agarose gel.

The choice of restriction enzyme to use in the digestion was limited, and no combination was found that would yield identifiable bands other than BamHI and XhoI. It was decided to prioritise the expression plasmid and to load a larger amount of DNA in order to attempt to see the smaller fragment corresponding to the DSB-inducing protein in each case. XhoI and BamHI digestion yields two fragments in the expression vectors: one 5912 bp fragment and one 480 bp (pQR1872), 1083 bp (pQR1873), 312 bp (pQR1874), or 441 bp fragment (pQR1875); digestion of pQR1871 yields 4242, 3000 and 1985 bp bands. The 5912 bp fragment of the expression vectors is very similar in size to the larger pQR1871 fragment, 4242 bp, and when loading a large amount of DNA, they are both undistinguishable. The gels can be observed in Figure 5-43 to Figure 5-47.
Figure 5-43. Miniprep digestions with XhoI and BamHI of DH5αZ1 pQR1871 pQR1872 (PvuII) cultures induced and grown overnight. Odd-numbered lanes are undigested controls, and even-numbered lanes are the digestions. Lanes 1 and 2 are a non-induced control; lanes 3 and 4 are induced replicate 1; lanes 5 and 6 are induced replicate 2; and lanes 7 and 8 are induced replicate 3. 2 µg of plasmid were loaded in each lane. The green square indicates where it is expected to observe the band corresponding to the 1985 bp fragment of pQR1871; the blue square indicates where it is expected to observe the band corresponding to the 480 bp fragment of the PvuII CDS in pQR1872. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).

Figure 5-44. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1873 (HpaII) cultures induced and grown overnight. Odd-numbered lanes are undigested controls, and even-numbered lanes are the digestions. Lanes 1 and 2 are a non-induced control; lanes 3 and 4 are induced replicate 1; lanes 5 and 6 are induced replicate 2; and lanes 7 and 8 are induced replicate 3. 2 µg of plasmid were loaded in each lane. The blue square indicates where it is expected to observe the band corresponding to the 1083 bp fragment of the HpaII CDS in pQR1873. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).
Figure 5-45. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1874 (ccdB) cultures induced and grown overnight. Odd-numbered lanes are undigested controls, and even-numbered lanes are the digestions. Lanes 1 and 2 are a non-induced control; lanes 3 and 4 are induced replicate 1; lanes 5 and 6 are induced replicate 2; and lanes 7 and 8 are induced replicate 3. 2 µg of plasmid were loaded in each lane. The blue square indicates where it is expected to observe the band corresponding to the 312 bp fragment of the ccdB CDS in pQR1874. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).

Figure 5-46. Miniprep digestions XhoI and BamHI of DH5αZ1 pQR1871 pQR1875 (denA) cultures induced and grown overnight. Odd-numbered lanes are undigested controls, and even-numbered lanes are the digestions. Lanes 1 and 2 are a non-induced control; lanes 3 and 4 are induced replicate 1; lanes 5 and 6 are induced replicate 2; and lanes 7 and 8 are induced replicate 3. 2 µg of plasmid were loaded in each lane. The blue square indicates where it is expected to observe the band corresponding to the 441 bp fragment of the denA CDS in pQR1875. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).
Figure 5-47. Miniprep digestions XhoI and BamHI of CSR603 pQR1871 pQR1872 (PvuII) cultures induced and grown overnight. Odd-numbered lanes are undigested controls, and even-numbered lanes are the digestions. Lanes 1 and 2 are a non-induced control; lanes 3 and 4 are induced replicate 1; lanes 5 and 6 are induced replicate 2; and lanes 7 and 8 are induced replicate 3. 2 µg of plasmid were loaded in each lane. The blue square indicates where it is expected to observe the band corresponding to the 480 bp fragment of the PvuII CDS in pQR1872. The ladder (M) is GeneRuler DNA Ladder Mix (ThermoFisher).

The results are not consistent for all strains. In the gel for DH5αZ1 pQR1871 pQR1872 digestions (Figure 5-43), the band corresponding to the 1985 bp fragment of pQR1871 cannot be observed for replicates 2 and 3 (green square). In addition, the smaller fragment corresponding to the PvuII CDS (480 bp, blue square) can only be seen slightly for replicate 1. In the case of DH5αZ1 pQR1871 pQR1873 (Figure 5-44), the smaller 1083 bp inserts corresponding to the HpaII CDS are present (blue square), although in the case of replicates 2 and 3 it migrates further in the gel (smaller blue squares). In DH5αZ1 pQR1871 pQR1874 (Figure 5-45), the 312 bp band corresponding to the ccdB CDS can be faintly seen for all samples (blue square). In the DH5αZ1 pQR1871 pQR1875 gel (Figure 5-46), the smaller 441 bp insert is clearly visible in the non-induced control, but not as much in the induced samples. Finally, in CSR603 pQR1871 pQR1872, the smaller 480 bp insert is present as expected (blue square).

The issue with these smaller inserts being visualised in the gels might have to do with their small size, and the fact that the expression plasmid is present in the miniprep in a lower proportion than pQR1871: this plasmid has an estimated 15-20 copies per cell, and the expression plasmid (pQR1872, pQR1873, pQR1874 or pQR1875), only an
estimated 5 copies per cell due to their different replication origins (ColE1 and pSC101 respectively), making the inserts belonging to the expression plasmid even more difficult to observe.

Minipreps were also sent to sequencing in order to determine whether there were additional mutations in the DSB-inducing protein CDS in the expression plasmid or in the exonuclease cassette in pQR1871. Oligonucleotides #S-03 and #S-04 were used for the expression plasmid, and #S-09 and #S-10 for pQR1871 (sequence available in the Appendix, section 8.2.1). The sequencing results indicated that there were no mutations other than the previously confirmed deletions in PvuII and ccdB (see section 5.2.2.2).

5.2.6.3 Investigating plasmid presence post-induction
Since the digestion pattern experiments are not conclusive, one last experiment was performed to investigate whether loss of activity could be related to plasmid loss. The different DSB-inducing protein expressing strains were induced and grown for a total of 24 h, then serial dilutions were performed and plated in triplicates on LB agar, LB agar with Amp, LB agar with Cam, and LB agar with both Amp and Cam. Colonies were counted and cfu / mL for each antibiotic plate was divided by the cfu / mL obtained in the LB agar control to quantify the proportion of cells that maintained the plasmid. These ratios are displayed in Table 5-2.
Table 5-2. Percentage of different E. coli strains expressing nucleases cells maintaining either the controller plasmid (Cam), the expression plasmid (Amp), or both (Amp + Cam) calculated taking the cfu / mL on LB agar with no antibiotics as the 100%.

<table>
<thead>
<tr>
<th></th>
<th>Amp</th>
<th>Cam</th>
<th>Amp + Cam</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5αZ1 pQR1871 pQR1872 (PvuII)</td>
<td>12.67 ± 1.15 %</td>
<td>0.006 ± 0.001 %</td>
<td>0.011 ± 0.006 %</td>
</tr>
<tr>
<td>DH5αZ1 pQR1871 pQR1873 (HpaI)</td>
<td>12.29 ± 1.76 %</td>
<td>0.005 ± 0.004 %</td>
<td>0.005 ± 0.002 %</td>
</tr>
<tr>
<td>DH5αZ1 pQR1871 pQR1874 (ccdB)</td>
<td>32.17 ± 4.26 %</td>
<td>1.06 ± 0.59 %</td>
<td>0.43 ± 0.094 %</td>
</tr>
<tr>
<td>DH5αZ1 pQR1871 pQR1875 (denA)</td>
<td>11.51 ± 2.08 %</td>
<td>0.017 ± 0.007 %</td>
<td>0.008 ± 0.03 %</td>
</tr>
<tr>
<td>CSR603 pQR1871 pQR1872 (PvuII)</td>
<td>121.9 ± 14.5 %</td>
<td>0.006 ± 0.001 %</td>
<td>0.004 ± 8 · 10⁻⁶ %</td>
</tr>
</tbody>
</table>

It appears that most cells lack pQR1871 (CamR) in all strains: only 0.005 – 1.06 % of cells retain it and grew on LB Cam agar. A greater proportion of cells retain the expression plasmid (Amp), although this is only true for DH5αZ1 strains: in the case of CSR603 pQR1871 pQR1872, despite most do not harbour pQR1871, the number of cfu / mL obtained when plating on LB Amp agar slightly surpassed the cfu / mL grown on LB agar alone (121.9 %, just over 100 % considering the standard deviation).

In addition, in the case of DH5αZ1 pQR1871 pQR1874 (expressing ccdB), a smaller proportion of the cell population harbours either plasmid: in this case, the proportion of cells retaining the expression plasmid (Amp) was 32.17 %, for the recombinase controller plasmid it is 1.06 % and 0.43 % for both respectively. This can be explained by the mode of action of these DSB-inducing proteins: PvuII, HpaI and denA are endonucleases, whereas ccdB is a gyrase poison that causes DNA damage indirectly by targeting gyrA, DNA gyrase subunit A, and stabilising the cleavage complex formed during its action. It is easier for cells to develop resistance to a DNA gyrase poison than against direct DNA cleavage, and a number of strains have developed mutations in gyrA that confers them resistance to ccdB; an example of such is E. coli DB3.1, which contains a gyrA462 allele.
These experiments suggest that DSB-inducing protein activity is being lost through plasmid loss, or plasmid-free cells outgrowing the rest of the population as the antibiotic degrades. Plasmid maintenance is a burden to cells, burden that is magnified by these encoding toxic proteins. In addition, the ability of CSR603 pQR1871 pQR1872 to grow on Amp as opposed to DH5αZ1 explains why the 506 bp insert corresponding to the PvuII CDS was much more evident in the miniprep digestions (see Figure 5-43 and Figure 5-47, blue squares). In order to increase plasmid stability, these plasmids could be engineered to include cer sites, or the expression system could be integrated in the host genome.
5.3 Summary and conclusions

In this chapter, several DBS-inducing proteins have been expressed together with exonucleases and evaluated as a potential genome ablation cassette. When cloning lethal proteins, if the expression system of choice does not completely repress expression, a small amount of undesired expression under non-induced conditions can either prevent successful cloning or promote outgrowth of cells harbouring mutations in the CDS. This is thought to be the reason why there were several difficulties cloning these DSB-inducing proteins: when using a high copy plasmid with a tac promoter, which is unable to completely repress expression in absence of inducer, few colonies were obtained and these harboured plasmids that had at least one mutation in the sequences for the restriction enzymes. The most common mutation was one-nucleotide deletions, causing a frame shift in the CDS and preventing expression of the correct protein. When a plasmid with a lower copy number was used instead, pQR445, not a single colony was obtained. A different approach involved first obtaining strains with functional methyltransferases, and then cloning the restriction enzymes in them. The methyltransferases of the NotI, PvuII, BamHI and HpaII restriction-modification systems were cloned into pACYC184 and their sequence confirmed, but when tested for in vivo activity via genomic DNA plug preparation and digestion, the E. coli genomic DNA was not methylated and was digested fully by the cognate enzymes. No explanation for this was found. A working PvuII methyltransferase was obtained when using the pPVU1 plasmid from P. vulgaris, but no compatible plasmid out of the available ones could be transformed in HB101 pPVU1t2.

The dual recombinase expression system eventually allowed DSB-inducing expression, although it was not reliable. When screening colonies, fragments of the wrong size were consistently obtained, and numerous colonies had to be screened until eventually the right size was amplified after colony PCR. It is unclear why this continued occurring. A possibility is that the presence of different att sites in the expression system may
somehow interfere with ligation. In addition, some preliminary considered positive colonies had no DSB-inducing activity upon induction when the plasmid was transformed in a strain containing the controller plasmid. This could be explained through the presence of undocumented cryptic promoters in the expression plasmid enabling a base level of expression, providing a selection pressure that enables outgrowth of mutants in the population.

More surprisingly, despite *PvuII* and *ccdB* were successfully cloned into the expression plasmid and their activity was confirmed, sequencing results revealed frame-shifting mutations in their CDSs. In the case of *PvuII*, an alternative start codon downstream of the mutation resulted in a truncated protein, and in the case of *ccdB*, the mutation resulted in a slightly smaller protein with 20 different amino acids at C-terminal. These unexpected results were not investigated any further because it deviated from the original aims of the project and being able to express functional DSB-inducing proteins was deemed sufficient to proceed.

The influence of induction time was also investigated. Escape frequencies were found to be much higher than the desired $10^{-8}$ (National Institutes of Health, 2016), and cell viability increases over time when inducing in the early exponential phase. Total DNA content estimations revealed that a significant proportion of DNA remains undigested in all conditions. Induction in the early exponential phase is more efficient, while induction in late exponential phase does not cause significant changes in DNA content compared to the non-induced control. The best results are obtained when inducing *PvuII* first and exonucleases later, and simultaneous induction early.

It is also discovered that escape frequency increases over time when inducing earlier in the exponential phase, whereas this does not occur when inducing later. Sequential induction does matter in terms of viability after 24 h. In addition, when inducing either only *PvuII* or exonucleases, the restriction enzyme is more lethal when produced earlier,
whereas exonucleases are more harmful when expressed later. Nevertheless, in
Chapter 2, it was revealed that addition of MMC and CI prevents escaper growth, so
DNA reduction is a more crucial metric than viability in this experiment. It was therefore
concluded that simultaneous induction early in the exponential phase yields the best
results.

Out of the four DSB-inducing proteins tested in DH5αZ1, *HpaII* was the best performing
in terms of DNA content and induced lethality. However, this protein could not be
expressed in CSR603. Comparison of CSR603 pQR1871 pQR1872 (*PvuII*) and DH5αZ1
pQR1871 pQR1873 (*HpaII*) determined that the former achieves greater reduction in
DNA content, and therefore is more suitable for development of a genome ablation
methodology.

When addressing the issue of an increased cell viability post-induction, it was found that
after overnight induction most cells were unable to grow on fresh antibiotic plates,
meaning that few were viable and maintaining the plasmid. Antibiotic degradation over
time allows outgrowth of plasmid-free cells, which eventually comprise most of the
population. This explains why there is no detectable restriction enzyme activity after 24 h,
and why only some DH5αZ1 pQR1871 pQR1873 cells can be re-induced when
inoculated again and produce *HpaII*. The source of plasmid DNA used in miniprep
digestions and sequencing reactions must be non-viable cells that still retain plasmid
DNA. In addition, the proportion of ccdB-expressing cells that grow on antibiotic is
significantly higher than endonuclease-expressing cells, which may be due to resistance
against DNA gyrase poisons through mutations in DNA gyrase subunit A (Bernard and
Couturier, 1992).

It is evident that genome ablation cannot be achieved through expression of DSB-
inducing proteins, alone or in combination with antibiotic addition. If plasmids were
engineered to include stabilising elements such as cer sequences, genome ablation
results could be improved. Inclusion of other types of DNA damaging agents, such as UV, may be the way forward, despite it is not an easily scalable method. This will be explored in the next chapter.
Chapter 6. Using UV radiation to achieve genome ablation and assessment of a genome-ablated chassis

6.1 Introduction

This work intends to achieve a genome ablation protocol that fulfils two requirements: firstly, to guarantee no cell survival in order to ensure that genome-ablated (GA) cells do not replicate in the case of release; and secondly, to ensure that GA cells contain no DNA that can be acquired by other microorganisms via mechanisms such as natural transformation or that can recombine with naturally present DNA. The past chapters have proven that expression of DSB-inducing proteins alone is not sufficient to achieve either of these requirements, but addition of MMC at 1 µg/mL has been proven to consistently inhibit *E. coli* CSR603 survival as seen on LB agar tests. UV light has not been evaluated yet in this work for its ability to induce DNA damage due to its poor scalability, but it has the potential to be a good alternative given its use as a Maxicell strain (Sancar et al., 1979).

In this chapter, *E. coli* CSR603 cultures harbouring plasmids pQR1871 and pQR1872 for *Pvu*II and exonuclease expression are subjected to different UV doses to determine which condition generates enough DNA photodamage. Once the optimal UV irradiation time is determined, this protocol is applied to CSR603 strains constitutively expressing different proteins that can be assayed easily to quantify the loss of protein activity: exonuclease III, T5 phage exonuclease and restriction enzyme *Pvu*II, induced as part of this protocol; superfolder GFP (sfGFP); a section of the *Pseudomonas* meta pathway for benzoate degradation; and the *Chromobacterium violaceum* CV2025 transaminase (TAm). GA cells are also evaluated as a recipient for transformation of an RFP plasmid via flow cytometry. Finally, GA cells are stained with a viability dye and a DNA-specific dye and analysed via flow cytometry and fluorescence microscopy to analyse the effect of the genome ablation protocol at a single cell level.
6.2 Results and discussion

6.2.1 Use of M9 to enhance reproducibility

In Chapter 3, section 3.2.4, and Chapter 4, 4.2.2, it was mentioned that in many instances when culturing *E. coli* CSR603 pQR1870 or CSR603 pQR1871 pQR1872, a decrease in OD$_{600}$ was not observed upon induction. When testing the induced cells for *in vitro* λ DNA degradation no significant digestion was observed, meaning that there was no detectable exonuclease or restriction enzyme production and the experiment had to be aborted.

The media of choice in this work was LB, a complex undefined media containing tryptone, yeast extract and sodium chloride commonly used in molecular biology. Tryptone is the source of amino acids in the media and is obtained through enzymatic hydrolysis of casein, a milk protein, meaning that LB may contain lactose, inducer of the *lac* promoter and thus the *tac* promoter, which is used to express exonuclease III and T5 exonuclease in this work. Some lactose can also be present in the yeast extract. This causes a basal level of expression in non-induced conditions that acts as a selective pressure, forcing cells to develop strategies to circumvent exonuclease expression. Since the exact composition of LB is unknown and varies batch to batch, this means that different batches will have more lactose than others, explaining the observed unpredictable behaviour of these *E. coli* strains. Addition of 1 % glucose to LB is often recommended to help repress this undesired induction of lac-derived promoters (Komai et al., 1997; Pan and Malcolm, 2000), but it did not help in this instance (data not shown).

Media selection is known to impact the performance of genetic logic gates and strain performance (Siegal-Gaskins et al., 2014). In contrast to LB, defined media have a known precise composition and may be a preferable alternative: supplemented minimal media such as M9 is frequently when working with genetic circuits (Bonnet et al., 2012; Chen et al., 2019; Ward and Lutkenhaus, 1985). Therefore, the use of M9 medium supplemented with 1 % glucose and 1 % casamino acids instead of LB was explored.
Casamino acids are obtained through acid hydrolysis of casein as indicated in the manufacturer’s instructions, instead of enzymatic hydrolysis like tryptone, resulting in a mixture composed of amino acids and some peptides and no lactose. Strain CSR603 pQR1871 pQR1872 was constructed again, but transformations were plated on M9 agar and cells grown on liquid M9 instead of LB. Five colonies were picked and grown to early exponential phase, transferred to a 96 well plate and induced, and the OD$_{600}$ was measured over 16 h. The growth curves are portrayed in Figure 6-1.
Figure 6-1. Growth curves of five different CSR603 pQR1871 pQR1872 colonies grown in M9 Amp Cam at different induction conditions: no induction (grey), L-arabinose at 1 % only (red), IPTG at 0.5 mM only (blue), or both L-arabinose and IPTG (green). 200 µL of each cell culture at the beginning of exponential phase was transferred in triplicates to 96 flat-well plates containing the appropriate inducer volumes and transferred to a TECAN Infinite 200 PRO plate reader at 37 °C. Absorbance at 600 nm was measured every 30 min over 16 h and plates were shaken before measurement. The error bars indicate the standard deviation.

Induction was followed by a clear decrease in OD_{600} for all five colonies, confirming that supplemented M9 media is a better option for CSR603 pQR1871 pQR1872 than LB. Therefore, all experiments from now on employ M9 liquid media and M9 agar plates for *E. coli* growth.
6.2.2 UV irradiation as an additional DSB-inducing agent

Chapters 4 and 5 demonstrated that genome ablation could not be achieved through nuclease expression alone or combined with antibiotic addition. A DSB-inducing agent that has not been explored yet in this work is UV light, which was commonly used in the Maxicell method in the 1970s and 80s due to its ability to create extensive chromosomal damage, enabling replication of plasmid DNA only (Sancar et al., 1979). UV irradiation induces pyrimidine adducts and cyclobutane-pyrimidine adducts that are skipped by DNA polymerase during replication, creating gaps in the strands to be repaired through recombination-dependent pathways (Smith, 2013).

E. coli CSR603 pQR1871 pQR1872 was grown and induced, and after 4 h UV-irradiated for different periods of time: 1, 5, 10, 15 and 20 min alongside a non-irradiated control. Details of this experiment can be found in Chapter 2, section 2.4.8. As explained in this section, the estimated UV fluence rate is 10 J / m² s. Irradiation experiments were performed by exposure time rather than by total fluence for easier experimentation, since fluence can be calculated from exposure time with this fluence rate. After irradiation, and similar to Maxicell protocols, cells were then resuspended in rich media and allowed to recover for 1 h so that viable cells escaping irradiation start to grow again, at which point ceftriaxone was added to target these escapers. Ceftriaxone is a cephalosporin that inhibits bacterial cell wall synthesis, therefore only affecting cells that are growing in this recovery stage (Hall et al., 1981). Modern minicell preparation protocols include ceftriaxone (Chen et al., 2019; Rampley et al., 2017), whereas older Maxicell protocols employ D-cycloserine instead, an antibiotic of similar mode of action (Heinemann and Ankenbauer, 1993; Jemiolo et al., 1988; Mayo et al., 1988). After an overnight incubation, 3 mL were plated on LB agar to detect escapers, and 5 mL were used to prepare DNA plugs in order to run FIGE as described in Chapter 2, section 2.2.1.12.

Quantification methods relying on DNA extraction were discarded in favour of FIGE, since it was found that some DNA is always lost in the DNA extraction procedure.
Comparisons not shown in this work revealed that, in some instances, when loading extracted DNA in an agarose gel no band could be seen, but when running FIGE with plugs prepared with the equivalent number of cells, a significant amount of DNA could be detected. Therefore, FIGE is the preferred method from now on. The FIGE gel obtained in this experiment can be observed in Figure 6-2.

Figure 6-2. FIGE image of UV-irradiated E. coli CSR603 pQR1871 pQR1872 (PvuII) cells irradiated at different times following the procedure explained in Chapter 2, section 2.4.8. Cells were either not induced (lanes 1-6) or induced at 1 % L-arabinose and 0.5 mM IPTG (lanes 7-12), and either not UV-irradiated (lanes 1, 7), or UV irradiated for 1 min (lanes 2, 8), 5 min (lanes 3, 9), 10 min (lanes 4, 10), 15 min (lanes 5, 11) or 20 min (lanes 6, 12). Lanes M are a HindIII-digested λ DNA marker.

Most genomic DNA is trapped within the plug and is not able to migrate (blue squares).

Nicked or partially open DNA can migrate a bit further (green squares), and these forms disappear after 1 min irradiation in both non-induced and induced samples.
As expected, the induced non-irradiated sample (lane 7) has less DNA than the induced control (lane 1) due to nuclease expression. No DNA is observed after 15 min of irradiation, but this occurs for both non-induced and induced samples. It is possible that there is a difference in DNA content at an individual cell level that cannot be observed in the gel, which will be addressed in section 6.3.2. In addition, no survivors were able to grow on LB agar in all samples that had been UV-irradiated, which guarantees that this method yields no cell viability.

This experiment was repeated with 15 min irradiation in triplicate, for both FIGE and determination of cell viability as the previous experiment. The FIGE gel is displayed in Figure 6-3.

Figure 6-3. FIGE gel of E. coli CSR603 pQR1871 pQR1872 (PvuII) induced samples, irradiated at 15 min (lanes 4-6) or not irradiated (lanes 1-3). The colours are inverted to ensure there is no DNA left. Lanes M are an EcoRV-digested λ DNA marker.
There is no visible DNA in any of the irradiated CSR603 pQR1871 pQR1872 samples, and no cells grew on LB agar. The faint smear observed in lanes 4-6 correspond to the agarose plug, not DNA, and is also observed in the marker lanes (M). 15 min of UV irradiation seems to satisfy the requirements of a genome ablation protocol: no cell viability, and elimination of DNA. Therefore, this protocol involving nuclease induction, UV irradiation for 15 min and ceftriaxone treatment is referred to as genome ablation from now on, and the cells prepared in this manner, genome-ablated cells (GA cells). Induction was still included in the protocol despite no difference was observed with the non-induced samples in order to add a layer of biocontainment.

In order to ascertain whether the resulting cells be used for their intended purpose, it is necessary to test whether proteins are still active. For this purpose, a series of plasmids were generated to assess whether there is still protein activity post-genome ablation.

### 6.2.3 Cloning constitutive expression cassettes into pQR1871

A constitutive protein expression strategy was designed in order to have cells producing the protein or proteins of interest prior to genome ablation. A cassette for constitutive protein expression was designed employing promoter J12119, the strongest of the Anderson constitutive promoter collection (Anderson, 2006); an RBS; a XhoI site; the sfGFP gene; a BglII site; and the rrnB T1 terminator. This was ordered as a gBlock, the sequence of which can be found in the Appendix, section 8.4.3, and cloned via Gibson assembly into the NcoI-EcoRI sites of pACYC184 (see Figure 6-4), yielding plasmid pQR1876 (see section 2.2.6.9). The plasmid map for pQR1876 can be found in the Appendix, section 8.1.23. However, CSR603 cells harbouring pQR1871, pQR1872 and pQR1876 grew very slowly. Instead, this cassette was cloned into the SmaI site of pQR1871 as explained in Chapter 2, section 2.2.6.10, and colonies screened through the presence of green fluorescence. This plasmid was termed pQR1877.
Figure 6-4. pACYC184 digestion with Ncol and EcoRI (lane 1) yielding fragments of 3944 bp (blue arrow) and 301 bp (green arrow, not clearly visible). The 3944 bp fragment was gel-purified. 500 ng of digested plasmid were loaded. The ladder (M) is the GeneRuler Mix DNA Ladder (Thermo Fisher).

It was then decided that different proteins were to be cloned into the XhoI-BglII site of pQR1877, between promoter and terminator: part of the Pseudomonas meta pathway; the C. violaceum CV2025 transaminase; and tinsel purple from ATUM. However, the portion of Pseudomonas meta pathway used here is a large 5 kb sequence and contained three XhoI sites and four BglII sites, which would complicate the cloning procedure via restriction digest. Therefore, the pQR1877 backbone was PCR-amplified to add a SmaI site and a BsrGI site instead, which are not present in the sequence for this pathway. The gel depicting this amplicon can be found in Figure 6-5.

Figure 6-5. PCR amplification of the pQR1877 backbone inserting SmaI and a BsrGI sites (lane 1, blue arrow, 10 kb). The bands in lane 1 at 2 kb and 8 kb are non-specific amplicons (orange arrows). The fragment was gel-purified. 1 µL of the PCR reaction were loaded. The ladder (M) is the 1 kb Extend DNA Ladder (NEB).
The different sequences were then PCR-amplified including *Sma*I and *Bsr*GI sites, digested and ligated into the pQR1877 backbone via a cycled ligation protocol (details on Chapter 2, section 2.2.1.8). Figure 6-6 depicts the PCR amplification of the catechol pathway sequence, and Figure 6-7, of tinsel purple and the CV2025 TAm. More details on the cloning procedure can be found in Chapter 2, section 2.2.6.10.

![Image](image1)

**Figure 6-6.** PCR amplification of the catechol pathway sequence (lane 1, blue arrow, 5229 bp) prior to digestion and ligation into the pQR1877 backbone. 1 µL of the PCR reaction was loaded. The ladder (M) is the GeneRuler Mix DNA Ladder (Thermo Fisher).

![Image](image2)

**Figure 6-7.** PCR amplification of the CV2025 TAm (lane 1, blue arrow, bp) and tinsel purple (lane 2, green arrow, bp) prior to digestion and ligation into the pQR1877 backbone. 1 µL of the PCR reaction was loaded. The ladder (M) is the 100 bp DNA Ladder (NEB).

The plasmid harbouring the fragment of the *Pseudomonas* meta pathway was termed pQR1878; tinsel purple, pQR1879; and the CV2025 TAm, pQR1880. Later on, it was decided to introduce a His-tag in the N-terminal of the TAm to enable detection via Western blotting. Since the His-tag was already present in the original template, pQR801 (Kaulmann et al., 2007), a PCR to obtain the fragment again and repeat the cloning
procedure was performed, but the PCR did not work despite repeated attempts probably due to the complexity of the repetitive sequence provided by inclusion of a His-tag. Instead, a gBlock spanning the beginning of the CV2025 TAm sequence within pQR801 flanked by *Sma*I and *Mau*BI sites was ordered (sequence available in the Appendix, section 8.4.4) and cloned via cycled ligation in *Sma*I and *Mau*BI-digested and purified pQR1880. Several colonies were picked and confirmed via sequencing using oligonucleotides #S-11 and #S-12 (sequence found in the Appendix, section 8.2.4). The resulting plasmid maps for all assemblies can be found in the Appendix, sections 8.1.24 to 8.1.28.

6.2.4 Assessing protein activity post-genome ablation

UV irradiation not only induces DNA damage: in proteins, UV light affects primarily aromatic amino acids, triggering electron ejection from their side chains (Bent and Hayon, 1975a, 1975b, 1975c) and generating covalent cross-links with other residues. Protein carbonylation also occurs through oxidation of amino acid side chains due to UV-generated ROS, damage that can be propagated across the whole protein structure (Hawkins and Davies, 2001; Krisko and Radman, 2013). Hence, it is necessary to determine whether GA cells can be a chassis for biocatalysis considering UV-induced photodamage. With this aim, *E. coli* harbouring pQR1872 (*Pvu*II) and pQR1871 or either of the plasmids mentioned in section 6.2.3 is subjected to UV irradiation and loss of activity is investigated for *Pvu*II and exonucleases (pQR1872), sfGFP (pQR1877), a portion of the *Pseudomonas* meta pathway (pQR1878), and the *C. violaceum* CV2025 transaminase (TAm, pQR1880).

Despite a plasmid for constitutive expression of tinsel purple was also obtained, pQR1879, it was not used because it only generated a faint purple colour that could not be quantified by doing a spectral scan in the spectrophotometer with empty CSR603 cells as blank.
6.2.4.1 *PvuII* and exonuclease activities

As proven in Chapter 5, *PvuII* activity is lost after inducing *E. coli* CSR603 pQR1871 pQR1872 and allowing it to incubate overnight due to plasmid loss or outgrowth of plasmid-free cells. *In vitro* activity assays were performed to ascertain if this different protocol allowed cells to maintain nuclease activity.

The same strain was induced and UV-irradiated at the same times as in section 6.2.2 and samples taken to test for restriction enzyme and exonuclease activities as explained in Chapter 2, section 2.3.2, including an additional 200 ng pUC19 and 500 ng λ DNA. Non-induced and induced samples were subjected to the UV irradiation protocol, together with controls consisting of samples taken 4 h post-induction. The gel depicting these in vitro assays can be observed in Figure 6-8.

![Figure 6-8. Gel depicting in vitro activity assays for *E. coli* CSR603 pQR1871 pQR1872 (*PvuII*) subjected to the UV irradiation protocol. Lanes are labelled in the image for easier interpretation. Lane A is an activity assay with pUC19 as substrate; lane B, with λ DNA as substrate; and lane C, with EcoRV-digested λ DNA as substrate. The first three lanes are control DNA with no protein extract addition. n/I denotes non-induced samples; 'Ind' denotes induced samples. 'n/I, 4 h' is a control assay performed with protein extract from growing *E. coli* cultures; 'Ind, 4 h' is the equivalent assay with a sample obtained 4 h after induction. The UV irradiation time is specified in the label for each set of assays.](image)

The non-induced sample 4 h post induction shows significant degradation of all three types of DNA corresponding to the background of endogenous cell nucleases. The induced control protein activity on pUC19 denotes that there is not much *PvuII* activity, since one clear band can be observed.
Looking at the induced samples subjected to the UV irradiation protocol, 1 min of UV irradiation does not have an effect on protein activity, whereas at 5 min of irradiation the impact is more noticeable and most protein activity is lost: in the EcoRV-digested λ DNA lane, the different fragments can be clearly seen. After 10 min irradiation the results clearly resemble those of the controls with no protein extract added to them, meaning that there is no nuclease activity.

When UV irradiation times were determined for efficiency of DNA degradation, it was shown that 10 min of exposure to UV started to cause a reduction in DNA content, where 15 min resulted in no DNA visible in the gel. Unfortunately, this genome ablation protocol implies loss of protein function, at least for nucleases. However, there might be proteins that are not affected by UV irradiation as much as these nucleases and different candidates are evaluated in the following sections.

6.2.4.2  sfGFP

sfGFP is a more stable version of the *Aequorea victoria* GFP engineered to be easily folded to enhance its applicability as a reporter via protein fusion (Pédelaq et al., 2006). Its robustness and ease of detection makes it an attractive candidate to evaluate presence of activity post-genome ablation. *E. coli* CSR603 pQR1877 pQR1872 was subjected to the UV irradiation protocol (see Chapter 2, section 2.4.8). A picture of the cell pellets under blue light can be found in Figure 6-9, and a quantification of the fluorescence units is depicted in Figure 6-10.
Figure 6-9. Pictures of the pellets of E. coli CSR603 pQR1877 pQR1872, non-induced (left) and induced (right), UV-irradiated at the times indicated above each tube followed by overnight incubation. In the picture, the cells are receiving blue light that is blocked from the objective with an orange screen. These same samples are quantified in Figure 6-10.

Figure 6-10. Graph depicting the relative fluorescence units (RFU) normalised by the absorbance at 600 nm of non-induced (light green) and induced (dark green) E. coli CSR603 pQR1872 pQR1877 subjected to different UV irradiation times and post-overnight incubation. Cells from Figure 6-9 were resuspended in PBS and 200 µL aliquots were added in triplicate to a 96 well plate, then transferred to a plate reader. Absorbance at 600 nm and green fluorescence (Ex488 / Em515) were measured. The error bars indicate the standard deviation.

The most notable difference between the two non-induced and induced samples is that a lower level of fluorescence is detected in the latter, even in absence of UV irradiation, indicating that nuclease expression halts transcription and translation to some extent. For both subsets, 1 and 5 min of exposure to UV causes an increase in fluorescence, while this decreases dramatically at 10 min, with little difference after this point. It is also
at the 10 min sample where the amount of DNA starts noticeably decreasing as seen via FIGE (see Figure 6-2), and where nuclease activity ceases to be observed as seen in in vitro assays (see Figure 6-8).

UV light is known to cause changes in gene expression. UV irradiation triggers the SOS response, a series of genes activated upon self-degradation of repressor LexA when ssDNA is detected by RecA. However, upregulation of other genes can occur via a LexA-independent mechanism, while UV irradiation can cause reduced transcription genes of others (Courcelle et al., 2001). It is possible that these transcriptional changes are upregulating expression of GFP under the J23119 promoter, explaining this increase in fluorescence at low UV irradiation times (1 and 5 min). In the higher UV doses, DNA damage must be too extensive for this to occur.

6.2.4.3 Pseudomonas meta pathway

*Pseudomonas* are Gram-negative bacteria commonly found in polluted environments such as contaminated water and soil and some species and strains are able to grow on xylenes and toluates, making then an attractive chassis for environmental biosensors (de Las Heras and de Lorenzo, 2012). The *Pseudomonas meta cleavage catabolic pathway* enables degradation of benzoate and methyl benzoates via 13 genes and is encoded in the *P. putida* plasmid pWWO-161. An intermediate of this pathway is 2-hydroxymuconic aldehyde, a bright yellow compound that can be monitored via absorbance measurement at 395 nm. A scheme depicting this pathway can be found in Figure 6-11.
This particular section of the meta pathway containing 6 genes was selected for evaluation of GA cells as chassis for whole metabolic pathways given the ease of 2-hydroxymuconic aldehyde detection via addition of sodium benzoate, the substrate of toluate 1,2-dioxygenase (xylXYZ), or addition of catechol, the substrate of catechol 2,3-dioxygenase (xylE). Therefore, this segment of the operon was cloned into plasmid pQR1871 yielding plasmid pQR1878 (as explained in Chapter 2, section 2.2.6.10), and this was electroporated into CSR603 pQR1872.

During the construction of strain CSR603 pQR1872 pQR1878, positive TOP10 colonies harbouring plasmid pQR1878 were picked and grown overnight on LB Cam in order to prepare plasmid DNA for electroporation into E. coli CSR603 pQR1872. Addition of sodium benzoate at 5 mM to 0.5 mL of this overnight culture resulted in 2-hydroxymuconic aldehyde production (Abs395 = 0.45), although it required an incubation at room temperature overnight instead of the instantaneous product generation observed when adding catechol at 5 mM. However, after electroporation of this plasmid into E. coli CSR603 pQR1872 and addition of sodium benzoate at 5 mM to the resulting CSR603 pQR1872 pQR1878 strain, no product was observed even after a 24 h incubation. It seems that this strain is not capable of converting benzoate into 2-hydroxymuconic aldehyde despite encoding the necessary enzymes. It does produce adequate amounts of catechol 2,3-dioxygenase, since addition of catechol at 5 mM results in an instant
yellow colouration. Thus, despite the aim of this experiment was to evaluate whether GA cells could perform as chassis for whole metabolic pathways, only the last enzyme, catechol 2,3-dioxygenase, was assayed via addition of catechol at 5 mM. The results can be observed in Figure 6-12.

When analysing the effect of UV irradiation on sfGFP, an increase in fluorescence was observed when irradiating for 30 s and 1 min followed by a decrease in both non-induced and induced subsets (see Figure 6-10). However, this is not observed here: no activity is observed even after 30 s irradiation. It is likely that the amount of sfGFP present when irradiating was already higher at the time of irradiation than of catechol 2,3-dioxygenase in this case: wild-type GFP has a half-life of approximately 26 h (Corish and Tyler-Smith, 1999), and it could be even longer for sfGFP. While the exact relationship between

Figure 6-12. Graph depicting the results of the catechol 2,3-dioxygenase assay on non-induced (light yellow) and induced (dark yellow) E. coli CSR603 pQR1872 pQR1878 subjected to different UV irradiation times and post-overnight incubation. Samples were resuspended in PBS and catechol was added at 5 mM. Samples were incubated for 15 min and transferred to 96 well plates. The experiment was done in triplicate. The blanks were the same sample with addition of an equivalent amount of PBS instead of catechol. Absorbance at 395 nm and 600 nm were measured with a TECAN Infinite M200 plate reader in order to normalise the amount of 2-hydroxymuconic aldehyde by cell density. The error bars indicate the standard deviation.
sfGFP structure and fluorescence emission is not well understood, it is a simpler structure than catechol 2,3-dioxygenase: sfGFP is a 28 kDa monomer, while catechol 2,3-dioxygenase is an oligomer consisting of four 31.5 kDa subunits with a ferrous iron each (Kita et al., 1999). It is no surprise then that catechol 2,3-dioxygenase appears to be more sensitive to UV photodamage and has a shorter half-life than sfGFP.

6.2.4.4  

C. violaceum CV2025 TAm

The *C. violaceum* CV2025 TAm is a highly stereoselective ω-aminotransferase with a wide amino acceptor substrate range, making it an attractive enzyme for amination in biotechnological applications. It is a homodimeric protein formed by two identical pyridoxal phosphate (PLP)-binding subunits of 51.2 kDa each. Both activity and presence of the CV2025 were investigated in this instance, since the latter can be easily verified through Western blotting of samples post-GA using a His-tag specific antibody. Plasmid pQR1886, which contains a His-tag in the N-terminal of the CV2025 TAm, was used in this experiment (see section 2.2.6.11).

Non-induced and induced *E. coli* CSR603 pQR1872 pQR1886 were irradiated at different times between 0 s and 15 min (0 s, 30 s, 60 s, 2 min, 5 min, 15 min). Samples were taken to perform the transaminase assay explained in Chapter 2, section 2.3.2.4, and to run a Western blot as explained in Chapter 2, section 2.3.4. The TAm assay results are portrayed in Table 6-1, the conditions of which are specified in Table 6-2, and the Western blot is depicted in Figure 6-13.
Table 6-1. Transaminase assay for UV-irradiated samples performed in triplicates alongside the controls. Row 1 is the complete assay; row 2 is the control minus amino donor; row 3 is the control minus amino acceptor; row 4 is the control minus protein extract. The exact composition of each assay is detailed in Table 6-2.

<table>
<thead>
<tr>
<th></th>
<th>No UV</th>
<th>30 s</th>
<th>1 min</th>
<th>2 min</th>
<th>5 min</th>
<th>15 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-induced</td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
</tr>
<tr>
<td>Induced</td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
<td><img src="Image" alt="Image" /></td>
</tr>
</tbody>
</table>
Table 6-2. Composition of the different conditions in the TAm assay depicted in Table 6-1. All assays were performed in a total volume of 200 µL in phosphate buffer at 100 mM (pH = 7.5).

<table>
<thead>
<tr>
<th>Row 1</th>
<th>Row 2</th>
<th>Row 3</th>
<th>Row 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 µL protein extract</td>
<td>50 µL protein extract</td>
<td>50 µL protein extract</td>
<td>-</td>
</tr>
<tr>
<td>10 mM benzaldehyde (amine acceptor)</td>
<td>10 mM benzaldehyde (amine acceptor)</td>
<td>-</td>
<td>10 mM benzaldehyde (amine acceptor)</td>
</tr>
<tr>
<td>25 mM 2-(4-nitrophenyl)ethan-1-amine hydrochloride (amine donor)</td>
<td>-</td>
<td>25 mM 2-(4-nitrophenyl)ethan-1-amine hydrochloride (amine donor)</td>
<td>25 mM 2-(4-nitrophenyl)ethan-1-amine hydrochloride (amine donor)</td>
</tr>
<tr>
<td>0.2 mM PLP</td>
<td>0.2 mM PLP</td>
<td>0.2 mM PLP</td>
<td>0.2 mM PLP</td>
</tr>
</tbody>
</table>
The product of the transamination reaction is observed as a red-brown precipitate at the bottom of the wells in Figure 6-12. In the control assay without acceptor added (benzaldehyde), an intense red colour can also be observed, which is due to aldehydes and ketones such as pyruvate present in the clarified lysate acting as acceptors. In the control assay without acceptor added (benzaldehyde), a red colour can normally be observed due to aldehydes and ketones such as pyruvate present in the clarified lysate acting as acceptors. The red observed in this experiment is unusually intense: the plates were incubated for a total of 48 h instead of the 18 h incubation period used by Baud et al. (2015) since this strain is not ideal for protein production and a longer incubation was needed to detect TAm activity, which may also lead to a higher conversion of background substrates. This is also manifested in the colony screening assay as faintly orange colonies (Baud et al., 2015) as was observed when performing the recombinant colony screening for pQR1880.

The lack of a quantitative analytical method makes it difficult to appreciate differences between non-induced and induced samples, but it seems that induced samples generate
a lower amount of product than their non-induced counterparts. After 5 min irradiation, a small amount of precipitate can be observed, but none can be observed at 15 min irradiation meaning that the TAm is no longer active. In addition, a very faint band is observed at that equivalent timepoint in the Western blot (Figure 6-13, lanes N6 and I6), meaning that the TAm itself has been extensively photodamaged and is no longer present, which explains the absence of product.

6.2.5 GA cells as recipients for transformation

So far, only sfGFP has been shown to still be active, retaining fluorescence, post-genome ablation. Another avenue to explore is to ascertain whether these cells can receive a plasmid through chemical transformation, and whether the cell machinery is capable of plasmid maintenance and transcription and translation of proteins encoded within that plasmid. GFP variants are often used as a reporter gene encoded within the plasmid DNA for transfection of mammalian cells to enable flow cytometry detection of fluorescent cells and transfection efficiency quantification (Ducrest et al., 2002; Hellweg et al., 2003). The development of more potent flow cytometers has also enabled transformant screening in bacteria with the same premise (Bennett et al., 2015).

In this experiment, GA CSR603 pQR1871 pQR1872 cells were subjected to the protocol for routine preparation of chemically competent cells detailed in Chapter 2, section 2.2.2, and then mixed with the Km\textsuperscript{R} RFP plasmid from the ATUM Protein Paint Box (ATUM, California, US) in a chemical transformation reaction as explained in section 2.2.4.

In this experiment, GA cells were transformed with an RFP plasmid and run in a flow cytometer in order to detect transformants. \textit{E. coli} CSR603 pQR1871 pQR1872 was genome-ablated and cells were subjected to the protocol detailed in Chapter 2, section 2.2.2 for routine preparation of chemically competent cells. Cells were then mixed with the Km\textsuperscript{R} RFP plasmid from the ATUM Protein Paint Box (ATUM, California, US) following the chemical transformation procedure detailed in section 2.2.4, with addition of
kanamycin to the SOC media and an overnight recovery step. Controls consisted of the same GA cells with no plasmid added, and regular chemically competent *E. coli* CSR603 with plasmid and without. 1 and 2 µg plasmid DNA were transformed in the regular competent cells control and the GA cells respectively to counter the low transformation efficiency observed consistently for strain CSR603. For flow cytometry analysis, RFP- and RFP+ gates were drawn by running CSR603 cells harbouring no plasmid, and CSR603 harbouring the RFP plasmid. The maximum number of events per run allowed by the instrument, 10 million, were recorded for each sample: heat-shock mediated bacterial transformation is an inefficient process, since it involves subjecting the cells to drastic temperature changes in order to destabilise the bacterial membrane and cells do not survive the heat shock or recover. The results for the flow cytometry experiments can be found in Figure 6-14 and Table 6-3.
Figure 6-14. Flow cytometry analysis of GA transformants with an RFP plasmid. RFP was excited with a 561 nm laser and emission was detected with a 585/29 filter. 50,000 events from the RFP- and RFP+ control samples and 10,000,000 from the transformation samples were collected. Blue dots represent RFP- events, and red dots represent RFP+ events. A depicts an overlaid histogram of RFP- and RFP+ cells used to draw the gates. B depicts the transformation of control CSR603 regular competent cells with 1 µg RFP plasmid, yielding 403 RFP+ events. C depicts the transformation of GA cells with 2 µg RFP plasmid yielding 3 RFP+ events.
Table 6-3. Flow cytometry results compilation of the RFP plasmid transformation experiment compiling data from the two RFP- and RFP+ controls, the regular competent and GA cells transformation controls with no plasmid DNA, and the regular competent and GA cells transformed with 1 and 2 µg DNA respectively. The frequency of RFP+ cells is the percentage of events within the total cell population. The number of potential transformants is indicated in bold in that row. MFI, or Median Fluorescence Intensity, is a measure of the shift in fluorescence used to compare different populations and is only applicable in subsets with RFP+ events.

<table>
<thead>
<tr>
<th></th>
<th>RFP- cells (control)</th>
<th>RFP+ cells (control)</th>
<th>control cells 0 µg plasmid</th>
<th>control cells 1 µg plasmid</th>
<th>GA cells 0 µg plasmid</th>
<th>GA cells 2 µg plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of RFP+ cells</td>
<td>0</td>
<td>99.9</td>
<td>0</td>
<td>4.23E-03 (403 events)</td>
<td>0</td>
<td>3.16E-05 (3 events)</td>
</tr>
<tr>
<td>MFI</td>
<td>-</td>
<td>106</td>
<td>-</td>
<td>29.3</td>
<td>-</td>
<td>110</td>
</tr>
</tbody>
</table>
As expected, no RFP+ events were detected in regular CSR603 and GA cells with no plasmid DNA. A total of 403 RFP+ events were detected in the CSR603 competent cell transformation with 1 µg plasmid DNA, whereas just 3 RFP+ events were detected in the GA cell transformation with 2 µg plasmid DNA. However, it cannot be guaranteed that the three events detected in the GA sample are indeed GA cells that have acquired the plasmid and are expressing RFP: it is possible that these events are bacteria expressing RFP that had been trapped in the fluidics of the system and not rinsed out between sample runs.

Interestingly, the median fluorescence intensity (MFI) of the GA RFP+ events is 110, similar to the 106 of control RFP+ cells. In contrast, regular transformed RFP+ events have an FMI of 29.3. If these GA RFP+ events are indeed transformants, this can be explained by an increased protein production from the RFP plasmid in a chromosome-free background, leading to a higher fluorescence per cell.

### 6.3.2 Flow cytometric analysis of GA cells

In the FIGE gel pictured in Figure 6-2, no obvious difference in DNA content was found between genomic DNA plugs from non-nuclease induced and nuclease-induced samples. These gels are convenient for a gross estimation of DNA content, more accurate than regular DNA extraction methods which always involve loss of sample. However, differences at a single cell level are not captured. A flow cytometry experiment was therefore designed to analyse GA cells and investigate the impact of nuclease expression and UV irradiation.

Cells were stained with two different dyes: a DNA dye, Hoechst 34580, and a viability dye, Live-or-Dye (488/515). Hoechst 34580 is a bisbenzimide, a cell-permeable DNA stain that binds to the minor groove of dsDNA, preferentially in AT-rich sequences (Weisblum and Haenssler, 1974), and Live-or-Dye 488/515 is a cell membrane impermeable amine-reactive dye that can only penetrate cells with damaged
membranes, discriminating between live (viable) or dead cells (non-viable). In the staining protocol detailed in Chapter 2, section 2.4.9, Live-or-Dye (LoD) is added prior to fixation to ensure that only dead cells are being detected. Four populations are expected: LoD- Hoechst-, the target population consisting of GA cells with a non-compromised membrane; LoD- Hoechst+, live cells containing DNA; LoD+ Hoechst-, dead cells with no DNA; and LoD+ Hoechst+, dead cells with DNA.

The samples were non-induced cells, either irradiated or non-irradiated; and induced cells, irradiated and non-irradiated. The live cells control was an E. coli CSR603 pQR1871 pQR1872 at the end of exponential phase. The dead cells control was obtained by resuspension of those cells in 70 % isopropanol. Unstained cells and fluorescence-minus-one (FMO) controls were also prepared and used in the analysis. The flow cytometric analysis can be found in Figure 6-15 and Table 6-4, and fluorescence microscope images of the samples can be found in Figure 6-16 and Figure 6-17.
Figure 6.15 Flow cytometry analysis of genome-ablated (GA) E. coli CSR603 pQR1871 pQR1872. Hoechst 34580 was excited with a 405 nm laser and emission was detected with a 520/35 filter; Live-or-Dye was excited with a 488 nm laser and emission was detected with a 542/27 filter. Hoechst 34580 fluorescence is portrayed in the X axis, and Live-or-Dye fluorescence in the Y axis. Gates were drawn using unstained and fluorescence-minus-one controls. The gated cells are depicted in red in the smaller FSC against SSC graph for each sample. The four samples are as follows: A: non-induced, non-UV irradiated; B: non-induced, UV-irradiated for 15 min; C: induced, non-UV irradiated; D: induced, UV-irradiated. A minimum of 50,000 events were collected for each sample, gated in the smaller graphs.
Table 6-4. Flow cytometry analysis of genome-ablated (GA) E. coli CSR603 pQR1871 pQR1872. The different quadrants in Figure 6-15 are referred to as LoD- or LoD+ and Hoechst- or Hoechst+ populations. The frequency is calculated with respect to the gated cells population (red events in the smaller graph in Figure 6-15). MFI, or Median Fluorescence Intensity, is a measure of the shift in fluorescence is only applicable in LoD+ and Hoechst+ subsets. The same data is included for the live and dead cells controls.

<table>
<thead>
<tr>
<th>Frequency of dead cells (LoD+)</th>
<th>Non-induced, non-irradiated</th>
<th>Non-induced, irradiated</th>
<th>Induced, non-irradiated</th>
<th>Induced, irradiated</th>
<th>Live cells control</th>
<th>Dead cells control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of LoD+, Hoechst- (Q1: dead, DNA-)</td>
<td>2.93</td>
<td>92.4</td>
<td>0.3</td>
<td>62.02</td>
<td>1.03</td>
<td>99.3</td>
</tr>
<tr>
<td>Frequency of LoD+ , Hoechst+ (Q2: dead, DNA+)</td>
<td>1.9</td>
<td>0.05</td>
<td>0.03</td>
<td>0.02</td>
<td>0</td>
<td>1.6 · 10^{-03}</td>
</tr>
<tr>
<td>Frequency of live cells (LoD-)</td>
<td>97.10</td>
<td>7.67</td>
<td>99.69</td>
<td>37.98</td>
<td>98.92</td>
<td>0.65</td>
</tr>
<tr>
<td>Frequency of LoD-, Hoechst+ (Q3: live, DNA+)</td>
<td>91</td>
<td>7.63</td>
<td>99.2</td>
<td>37.9</td>
<td>98.9</td>
<td>0.52</td>
</tr>
<tr>
<td>Frequency of LoD-, Hoechst- (Q4: live, DNA-)</td>
<td>6.1</td>
<td>0.04</td>
<td>0.49</td>
<td>0.08</td>
<td>0.02</td>
<td>0.13</td>
</tr>
<tr>
<td>Frequency of Hoechst+ cells</td>
<td>92.0</td>
<td>99.9</td>
<td>99.5</td>
<td>99.9</td>
<td>99.9</td>
<td>99.8</td>
</tr>
<tr>
<td>MFI Live-or-Dye</td>
<td>7.97</td>
<td>31.8</td>
<td>10.7</td>
<td>22.4</td>
<td>29.4</td>
<td>33.7</td>
</tr>
<tr>
<td>MFI Hoechst 34580</td>
<td>19.5</td>
<td>27.6</td>
<td>13.9</td>
<td>21.9</td>
<td>24.6</td>
<td>19.0</td>
</tr>
</tbody>
</table>
Figure 6-16. Widefield microscope images of non-induced E. coli CSR603 pQR1871 pQR1872 cells. A are images of the non-UV irradiated cells; B are images of the cells UV-irradiated for 15 min. 1 is the green channel only, 2 is the blue channel only, and 3 is the overlaid brightfield, blue and green channels. The white scale bar represents 10 µm.
Figure 6-17. Widefield microscope images of induced E. coli CSR603 pOR1871 pOR1872 cells. C are images of the non-UV irradiated cells; D are images of the cells UV-irradiated for 15 min. 1 is the green channel only, 2 is the blue channel only, and 3 is the overlaid brightfield, blue and green channels. The white scale bar represents 10 µm.
The aim of this experiment was to examine the impact of nuclease induction and UV irradiation on DNA content at an individual cell level through the differential binding of a DNA-reactive dye (Hoechst 34580, blue fluorescence), as well as to relate this to changes in cell viability as measured by the proportion of cells with damaged membranes that can be penetrated by an amine-reactive dye (Live-or-Dye, green fluorescence). As observed in section 6.2.2, FIGE analysis of genomic DNA plugs of induced and irradiated CSR603 pQR1871 pQR1872 reveal a decreased DNA content overall in induced samples and no DNA after 15 min of UV irradiation (see Figure 6-2 and Figure 6-3). Therefore, the expected result of the flow cytometry analysis would be a smaller population of blue events and a lower median of fluorescence intensity (MFI) for Hoechst+ events in induced and irradiated cells. However, this is not the case: the frequency of Hoechst+ cells in induced and UV-irradiated samples is 99.5 % - 99.9 % of all events, while in non-induced non-irradiated samples it is 92 % (see figure A in Figure 6-15, and A1-A3 in Figure 6-16). In addition, irradiated samples have higher MFI values for Hoechst+ events than non-irradiated: 27.6 versus 19.5 in the non-induced, and 21.9 versus 13.9 in the induced (see last row in Table 6-4). This is also manifested in the fluorescence microscopy images in Figure 6-16 and Figure 6-17, where a different staining pattern can be also observed between non-irradiated (A2 in Figure 6-16 and C2 in Figure 6-17) and irradiated samples (B2 in Figure 6-16 and D2 in Figure 6-17): UV-irradiated cells are emitting blue fluorescence more uniformly and brightly than their non-irradiated counterparts. Nevertheless, when comparing the four samples in terms of non-irradiated and irradiated subsets, MFI value is lower in both cases in the induced cells than the non-induced (13.9 versus 19.5 and 21.9 versus 19.5 respectively), a tendency that aligns more with the expectations of this experiment.

Hoechst 34580 is a cell-permeable dye that is often used to stain nuclei to identify cells in microscopy imaging (Bucevičius et al., 2018; Jäger et al., 2015; Na et al., 2015; Nogueira et al., 2016; Stowe et al., 2015). Hoechst stains and other DNA-binding dyes
are also used as an indicator of cell ploidy in cell cycle analysis (Gomes et al., 2018; Stowe et al., 2015). No attempts to monitor DNA degradation in vivo by endogenously expressed nucleases or UV irradiation through DNA-binding dyes have been reported in the literature, and when employing nucleases as a kill switch the emphasis is often on the effect on cell viability and not DNA degradation (Caliando and Voigt, 2015; Elešnik et al., 2016; Torres et al., 2003). The observed behaviour of Hoechst staining in this experiment could only be explained by an increased fluorescence emission when binding to mostly degraded and crosslinked DNA than to regular dsDNA, but this has not been discussed elsewhere.

The Live-or-Dye stain does behave as would be expected, as demonstrated by the controls: 99.3 % of events in the dead cells control display green fluorescence due to their damaged membranes, and only 1.03 % of events in the live cells control is stained (see Table 6-4). UV irradiation is causing a significant loss of viability: the frequency of LoD+ events (dead cells) increases from 2.93 % to 92.4 % in the non-induced subset, from 0.3 % to 62.02 % in the induced samples. However, when looking at the UV-irradiated subset, the frequency of viable or ‘live’ cells is higher in the induced samples (37.98 %) than in the non-induced 7.67 %). This also happens to some extent in the non-irradiated subset (99.69 % versus 97.1 %). In the flow cytometry analysis in Figure 6-15, this is manifested by the shift of the cell population in the induced samples (figures C and D in Figure 6-15) towards the LoD- quadrants with respect to their non-induced equivalents (Figure 6-15 A and B respectively), and a new distinct population can even be observed in the induced and irradiated cells (Figure 6-15 D, Q3). This implies that nuclease induction causes a decrease in lethality in the context of this genome ablation protocol.
6.3 Conclusions

In this chapter, the aim of achieving a genome ablation strategy appears to have been fulfilled. UV irradiation at 15 min and ceftriaxone treatment achieves both objectives of a genome ablation strategy: to ensure no cell survival, and to eliminate genomic and plasmid DNA as measured by FIGE in order to obtain a completely orthogonal chassis that will not interfere with the environment via replication or transmission of DNA. Ceftriaxone may contribute to GA cell production, but it is not an integral part of the protocol since its function is to ensure cells escaping UV are destroyed, similar to minicell production protocols. In addition, the shift to M9 media instead of LB enabled more reproducible results and has thus been proven to be a better alternative when working with genetic switches.

FIGE gels revealed no DNA remaining after 15 min of UV irradiation, but no difference in DNA content was observed between nuclease-expressing and non-induced samples. The protocol for GA cell production including induction of these nucleases was followed regardless in order to add an additional layer of biocontainment. Different proteins of varying structure and modes of action were constitutively expressed prior to induction and UV irradiation. *PvuII* activity is not observed clearly even in the control and is probably lost completely over the course of the genome ablation protocol, which lasts for just over 24 h after inoculation. Exonuclease activity starts to be lost at 5 min of UV irradiation and ceases to be measurable at 10 min of irradiation. sfGFP fluorescence was still visible to the naked eye even after 20 min of irradiation, albeit the measured fluorescence output was 40% that of the non-irradiated sample in induced samples and just over half that of the non-irradiated in non-induced samples. Catechol 2,3-dioxygenase activity is lost at just 30 s of irradiation, which is likely due to its complex oligomeric iron-binding structure. Induced samples also display a lower activity, which may be due to induction causing plasmid-free cells to outgrow the rest, leading to a lower amount of protein produced. This is also observed in the case of the CV2025 TAm,
although in this case only a qualitative estimation of TAm activity was possible. Here, while some activity is preserved after 5 min irradiation, this is completely lost at 15 min of UV irradiation. These results indicate that there is a trade-off between DNA content and metabolic activity: the UV radiation dose required for genome ablation does not allow for protein activity.

GA cells were also evaluated as a recipient for transformation in an experiment in which an RFP plasmid was transformed and cells were analysed via flow cytometry. 403 events were detected in a control consisting of chemically competent *E. coli* CSR603 cells, which represented a frequency of $4.23 \cdot 10^{-3} \%$, much less efficient than frequencies of approximately 0.1 % reported in other works (Bennett et al., 2015). This is probably due to the fact that *E. coli* CSR603, as mentioned in the previous chapters, is a weak, slow-growing strain not best suited for transformation; in fact, fewer transformants were consistently obtained throughout this project when using this strain than when using DH5αZ1, TOP10 or HB101. Transformation of GA cells yielded only 3 events, which represent a frequency of $3.16 \cdot 10^{-5} \%$. It is possible that these events are not actual transformants, but RFP+ cells ran as the control that remained in the fluidics of the flow cytometer. A total of 10,000,000 events were collected to increase the probability of finding transformants, but this also means that it is more likely that these cells trapped in the fluidics would be detected. These cells would have to be sorted and analysed to ensure that they were real transformants, but time limitations did not allow for this analysis.

Staining and visualisation of GA cells did not have the expected results. Cells were stained with a viability dye, Live-or-Dye, and a cell-permeable DNA-binding dye, Hoechst 34580, and analysed via fluorescence microscopy and flow cytometry. Despite FIGE gels had previously revealed that 15 min of UV irradiation caused enough DNA photodamage to the extent that no DNA would be visible in the gel, the vast majority of cells were successfully stained with Hoechst 34580 and the irradiated samples displayed
a higher Hoechst MFI. However, in the non-irradiated subset, induced cells were less fluorescent than their non-induced counterparts, suggesting that there is indeed a lower DNA content. The hypothesis is that Hoechst 34580 has a different mode of action with degraded, cross-linked DNA than with regular dsDNA, which is its usual application. No other works in the literature have been found to employ DNA-binding dyes in this manner. Other dyes such as DAPI and Hoechst 33342 could not be used since the flow cytometer lacked a UV laser: despite manufacturers report that these dyes can be excited with violet lasers, trial runs failed to detect any fluorescence with the 405 nm laser and eventually Hoechst 34580 was used. Nevertheless, viability staining reveals that the frequency of non-viable cells is reduced when inducing nucleases in both non-irradiated and irradiated samples. Therefore, this suggests that the aim of achieving a working GA cell would be possible if the mechanism of DNA-damaging protein induction was optimised for both DNA degradation and inhibition of cell recovery.
7 Chapter 7. Conclusions and future work

7.1 Summary and conclusions

This work proposes genome-ablated cells as a novel method of biocontainment and relays the requirements of an effective genome ablation strategy: to achieve absolute DNA degradation in vivo and prevent cell replication whilst preserving protein function and metabolic activity.

The first component of a genome ablation platform is exonuclease overexpression. Simultaneous expression of exonuclease III and T5 exonuclease was found to be the most efficient combination, achieving the lowest DNA content at 4 h post-induction. Examination of cell morphology via brightfield microscopy reveals that exonuclease expression does not induce cell filamentation, which normally occurs upon extensive DNA damage.

The second component of this platform is the DNA-cleaving agent, of which several were examined: antibiotics mitomycin C and ciprofloxacin, expression of double strand break-inducing proteins, and UV irradiation. Addition of mitomycin C at 1 µg / mL together with ciprofloxacin ensured no cell survival in E. coli CSR603 and DH5αZ1, but DNA remained at all concentrations assayed. The most efficient condition was CSR603 overexpressing exonucleases with addition of mitomycin C at 1 µg / mL and ciprofloxacin at 200 µg / mL 4 h post-induction. Additional induction of restriction enzyme PvuII yielded a similar extent of DNA degradation. Nevertheless, it is possible that this remaining DNA could be too damaged to be recombined with other DNA or transferred to other organisms.

Double strand break-inducing proteins were assessed next, but cloning took numerous attempts, including cloning under a tac promoter in a high (pTTQ18) and medium copy (pQR445) vectors, cloning under the araBAD or rha promoters, and cloning restriction enzymes in a strain with their cognate methyltransferase. A dual plasmid system for recombinase-mediated expression was eventually successful, and even then, the range
of proteins successfully expressed was limited: restriction enzymes *Pvu*II and *Hpa*II, gyrase poison *ccdB*, and endonuclease *denA*. Experimental work determined that simultaneous induction of *Pvu*II and exonucleases in *E. coli* CSR603 achieved the lowest DNA content, but none of the conditions tested resulted in escape frequencies lower than $10^{-8}$ (National Institutes of Health, 2016), and an increase in cell viability with prolonged culture growth was consistently observed after induction together with a loss of endonuclease activity.

UV irradiation for 15 min proved to be the optimal agent, preventing cell viability and achieving absolute DNA degradation. The established genome ablation protocol consisted on inducing *Pvu*II and exonucleases in *E. coli* CSR603 at the early exponential phase of growth, UV irradiation for 15 min after 4 h, and overnight incubation with ciprofloxacin. Evaluation of genome-ablated cells as a chassis for biocatalysis determined that there is a trade-off in UV dose and extent of photodamage, precluding biocatalysis. Assessment of genome-ablated cells as recipients of bacterial transformation was inconclusive, and while three positives were obtained these may be artefacts. Flow cytometric analysis and fluorescence microscope observation of genome-ablated cells revealed a staining pattern contradictory with FIGE results, suggesting that Hoechst 34580 behaves differently with endogenously degraded DNA, which has not been reported elsewhere. In addition, viability staining revealed that UV irradiation compromises cell integrity, which is detrimental to a functional chassis.

While the ambitious objective of achieving a completely functional genome-ablated chassis has not been fulfilled, this work highlights the considerations to take when expressing lethal DNA-damaging proteins and the unintended effects of other DNA-damaging agents. In addition, it highlights the limitations of analytical methods to quantify a reduced, or absence of, DNA content *in vivo* in order to obtain a DNA-free chassis, an application not reported before in the literature.
7.2 Future work

Further experiments could be performed to complement this work. The substitution of LB broth for supplemented M9 media was introduced rather late, but it enabled a consistent behaviour of \textit{E. coli} CSR603 upon nuclease induction. As mentioned throughout this thesis, a deceleration in growth was sometimes not observed when growing on LB broth, no nuclease activity was detected, and the experiment was aborted. This was solved with the use of supplemented M9 media. It would be interesting to repeat several experiments with supplemented M9 media instead of LB, since a tac promoter is involved in most of this work for exonuclease overexpression and the absence of lactose would imply a lower selective pressure pre-induction, potentially leading to an increased induction ratio and enhanced exonuclease activity. In addition, M9 media may provide a more defined environment and improve the cloning procedure of DSB-inducing proteins described in Chapter 5, leading to an easier and successful cloning of a larger range of toxic proteins with more straightforward screening. Further screening of tighter expression systems may determine a more suitable candidate for DNA damaging proteins.

The discussion on Chapter 4 suggests that, despite addition of MMC and CI together with exonuclease overexpression does not achieve absolute DNA degradation, this DNA may not be able to be transferred to other organisms or recombine with external DNA. This could be addressed with additional experiments. Plasmid DNA could be extracted from the nuclease-expressing and antibiotic-treated cells, and a sample transformed into competent \textit{E. coli} and plated on selective media to observe the appearance of transformants. In addition, an \textit{in vitro} recombination experiment could determine whether genetic information could be transferred in this manner.

Loss of nuclease activity observed in Chapter 5 was attributed to plasmid loss over time. Addition of cer sequences, which prevent plasmid multimerisation and loss, may circumvent this and enhance DNA degradation through higher levels of nuclease
expression. In addition, genome integration may also be contemplated: despite the constituents of complex synthetic networks are usually encoded in different plasmids with variable copy number in order to ensure each component is being synthesised at the appropriate quantities (Brophy and Voigt, 2014), this approach could be feasible if the necessary copies of each cassette are integrated into the E. coli genome.

Evaluation of GA cells as chassis for biocatalysis in Chapter 6 determined that the dose of UV irradiation needed for absolute DNA degradation prevented the biocatalytic function of catechol 2,3-dioxygenase and the CV2025 TAm, two oligomeric proteins. Given that sfGFP, a monomer, was still able to fluoresce at that UV dose, it is possible that other proteins with a simpler structure could survive the treatment and maintain their biocatalytic activity. In addition, when evaluating GA cells as recipients for bacterial transformation by flow cytometry, the transformants could be cell-sorted and incubated on LB for several days to ascertain whether these are true GA cells, or artefacts due to cells having escaped the treatment or trapped in the fluidics of the equipment. In addition, this method of UV irradiation only allows for genome ablation of 10 mL of cells at an OD = 0.2. It would therefore also be interesting to explore the possibility of scaling-up this method by using UV irradiation systems such as those used for viral inactivation in bioprocessing.

With regards to the inconclusive results of the DNA content analysis via flow cytometry, Hoechst 34580 was used due to the limitations of the system, which did not have a UV laser. DNA dyes such as DAPI or others from the Hoechst family could have been used instead if there had been access to a different flow cytometer with more excitation lasers and filter configurations.

With these considerations, subsequent projects could improve upon this work and eventually obtain a functional, metabolically active genome-ablated chassis. This chassis could also be assessed as recipient of bacterial conjugation, which has been
performed in minicells (Kobayashi, 2018), and even considered for genome transplantation (Hutchison et al., 2016; Lartigue et al., 2007).
8 Appendix

8.1 Plasmid maps

A map for each plasmid used as backbone and assembled in this work is appended in this section.

8.1.1 pTTQ18

Figure 8-1. pTTQ18 plasmid map extracted from SnapGene.
8.1.2 pQR445

Figure 8-2. pQR445 plasmid map extracted from SnapGene.
8.1.3 pACYC184

Figure 8-3. pACYC184 plasmid map extracted from SnapGene.
8.1.4 pQR1866

Figure 8-4. pQR1866 map extracted from SnapGene. A simplified version can be found in Chapter 3, Figure 3-1.
Figure 8-5. pQR1867 map extracted from SnapGene. A simplified version can be found in Chapter 3, Figure 3-1.
Figure 8-6. pQR1868 map extracted from SnapGene. A simplified version can be found in Chapter 3, Figure 3-1.
Figure 8-7. pQR1869 map extracted from SnapGene. A simplified version can be found in Chapter 3, Figure 3-1.
8.1.8 pQR1870

Figure 8-8. pQR1870 map extracted from SnapGene. A simplified version can be found in Chapter 3, Figure 3-1.
Figure 8-9. pACYC184-NotIM map extracted from SnapGene.
Figure 8-10. pACYC184-PvuIIM map extracted from SnapGene.
Figure 8-11. pACYC184-BamHIM map extracted from SnapGene.
8.1.12 pACYC184-HpaIIM

![Diagram of pACYC184-HpaIIM map extracted from SnapGene.]

Figure 8-12. pACYC184-HpaIIM map extracted from SnapGene.
Figure 8.13. pPVU1 map extracted from SnapGene.
8.1.14 pPVU1t

Figure 8-14. pPVU1t map extracted from SnapGene.
8.1.15 pPVU1t2

Figure 8-15. pPVU1t2 map extracted from SnapGene.
8.1.16 Controller plasmid from the dual recombinase expression system

Figure 8-16. Controller plasmid map extracted from SnapGene.
8.1.17 Expression plasmid from the dual recombinase expression system

Figure 8-17. Expression plasmid map extracted from SnapGene.
Figure 8.18. pQR1871 map extracted from SnapGene.
Figure 8-19. pQR1872 map extracted from SnapGene.
Figure 8.20. pQR1873 map extracted from SnapGene.
Figure 8-21. pQR1874 map extracted from SnapGene.
Figure 8-22. pQR1875 map extracted from SnapGene.
8.1.23 pQR1876

Figure 8-23. pQR1876 map extracted from SnapGene.
8.1.24 pQR1877

Figure 8-24. pQR1877 map extracted from SnapGene.
Figure 8-25. pQR1878 map extracted from SnapGene.
Figure 8-26. pQRR1879 map extracted from SnapGene.
8.1.27 pQR1880

Figure 8-27. pQR1880 map extracted from SnapGene.
8.1.28 pQR1886

Figure 8-28. pQR1886 map extracted from SnapGene.
### 8.2 Oligonucleotide list

#### 8.2.1 Sequencing oligonucleotides

*Table 8-1. Oligonucleotides used for sequencing using Eurofins Genomics tube sequencing service (Ebersperg, Germany).*

<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>M13 uni (-49)</td>
<td>GAGCGGATAACAATTTTACACACAGG</td>
</tr>
<tr>
<td>M13 uni (-21)</td>
<td>TGAAAACGACGGCCAGT</td>
</tr>
<tr>
<td>S-01</td>
<td>CATGGAAGCCATCACAG</td>
</tr>
<tr>
<td>S-02</td>
<td>CGTTCAGCTGGATATTACG</td>
</tr>
<tr>
<td>S-03</td>
<td>GAAGGTGAGCCAGTG</td>
</tr>
<tr>
<td>S-04</td>
<td>GTCTATAATCAGCGCAG</td>
</tr>
<tr>
<td>S-05</td>
<td>GAGAGGCTTCCAGCGC</td>
</tr>
<tr>
<td>S-06</td>
<td>CGTCGTCGACAGCAG</td>
</tr>
<tr>
<td>S-07</td>
<td>GTCTCGTTATCTACAG</td>
</tr>
<tr>
<td>S-08</td>
<td>GATTCATGTGTTGTTGAG</td>
</tr>
<tr>
<td>S-09</td>
<td>CTCCCGTTCTGGATAAT</td>
</tr>
<tr>
<td>S-10</td>
<td>CTCTCGCTATTACGCC</td>
</tr>
<tr>
<td>S-11</td>
<td>CTGCCAGTCTGCTACG</td>
</tr>
<tr>
<td>S-12</td>
<td>ATCCCACAGGTAGACG</td>
</tr>
</tbody>
</table>
### 8.2.2 Oligonucleotides used in Chapter 3

Table 8-2. Oligonucleotides used for gene amplification in cloning in Chapter 3. For oligonucleotides with highlighted nucleotides: the nucleotides in bold black are the overlapping sequences with the parental vector; the nucleotides in italics are the restriction site; the nucleotides in red correspond to a stop codon introduced to interrupt the lacZα CDS; the nucleotides in bold light blue are the RBS; the nucleotides in regular blue are random spacers; the underlined nucleotides are the sequences homologous to the gene or genes.

<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Sequence (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-01</td>
<td>CAGCGATGAATCCAGCTCGGTAATTAGGAGGAAATCGATATGAAATTTGTCTTTTTAATATCAACGGCCTG</td>
</tr>
<tr>
<td>1-02</td>
<td>GCAGGTCGACTCTAGAGGATCCCCTAGCGGCGGAAGGTGC</td>
</tr>
<tr>
<td>1-03</td>
<td>CAGCGATGAATCCAGCTCGGTAATTAGGAGGAAATCGATATATGACAAGGGACATTATCCTGCAGC</td>
</tr>
<tr>
<td>1-04</td>
<td>GCAGGTCGACTCTAGAGGATCCCTAGCGGCGGAAGGTCG</td>
</tr>
<tr>
<td>1-05</td>
<td>CAGCGATGAATCCAGCTCGGTAATTAGGAGGAAATCGATATGAGTAAATCCTGGGGAAAATTTATTGAAGAAG</td>
</tr>
<tr>
<td>1-06</td>
<td>GCAGGTCGACTCTAGAGGATCCCTAGCGGCGGAAGGTCG</td>
</tr>
<tr>
<td>1-07</td>
<td>GCGCCCAATACGGAAAAC</td>
</tr>
<tr>
<td>1-08</td>
<td>CCTCTTGCTATTACGCCAGC</td>
</tr>
<tr>
<td>1-09</td>
<td>GGTGTCATGATCGCTCTCTTTTAGCGGCGGAAGGTGC</td>
</tr>
<tr>
<td>1-10</td>
<td>GATCGCTCTCTTTTAGCGGCGGAAGGTGC</td>
</tr>
<tr>
<td>1-11</td>
<td>CCGCCAAGAGAGCGATCATGACACGGGACATTATCCTGCAGC</td>
</tr>
<tr>
<td>1-12</td>
<td>CCGCCAAGAGAGCGATCATGAAATCTGCGGAAAATTTATTGAAGAAG</td>
</tr>
</tbody>
</table>
### 8.2.3 Oligonucleotides used in Chapter 5

Table 8-3. Oligonucleotides used for cloning in chapter 5. For oligonucleotides with highlighted nucleotides: the nucleotides in bold black are the overlapping sequences with the parental vector; the nucleotides in italics are the restriction site; the nucleotides in bold blue are the RBS; the nucleotides in regular blue are random spacers; the nucleotides in bold orange are the restriction sites; the underlined nucleotides are the sequences homologous to the gene or genes; and the regular letters are additional nucleotides added to allow restriction enzyme binding.

<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Sequence (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-01</td>
<td>GTGAGCGGATAACAATTTTCACACAGAAACAGAATAGGAAGATCGATTAGAGCTCACTGCTCAAAATAAATTAGAGCTTTTG</td>
</tr>
<tr>
<td>2-02</td>
<td>CAGGTAGCAGATCCCCGGGTTAGTAAATCTTTGTCCATGGTTCCATTACATAATTAN</td>
</tr>
<tr>
<td>2-03</td>
<td>GACTGACATGGAATTCGTAATTTCTCTATGTGACAGCTTTATCATCGATAAGC</td>
</tr>
<tr>
<td>2-04</td>
<td>TACGCATGGAATTCTCACAGTCGGATGGCAGCCG</td>
</tr>
<tr>
<td>2-05</td>
<td>CAGCGATGAACTAGCTCGGTACCCTAGAGGAAATCGATATCGGAGTCG</td>
</tr>
<tr>
<td>2-06</td>
<td>GCAGGTGCACTCTAGAGGATCCCCCATGCACTCAAAACTTT</td>
</tr>
<tr>
<td>2-07</td>
<td>CAGCGATGAACTAGCTCGGTACCCTAGAGGAAATCGATATCGGAGTCG</td>
</tr>
<tr>
<td>2-08</td>
<td>GCAGGTGCACTCTAGAGGATCCCCCATGCACTCAAAACTTT</td>
</tr>
<tr>
<td>2-09</td>
<td>CAGCGATGAACTAGCTCGGTACCCTAGAGGAAATCGATATCGGAGTCG</td>
</tr>
<tr>
<td>2-10</td>
<td>GCAGGTGCACTCTAGAGGATCCCCCATGCACTCAAAACTTT</td>
</tr>
<tr>
<td>2-11</td>
<td>CAGCGATGAACTAGCTCGGTACCCTAGAGGAAATCGATATCGGAGTCG</td>
</tr>
<tr>
<td>2-12</td>
<td>GCAGGTGCACTCTAGAGGATCCCCCATGCACTCAAAACTTT</td>
</tr>
<tr>
<td>2-13</td>
<td>GACATCTAGCAAGCTCGCCGTTCCATCATAGGATTT</td>
</tr>
<tr>
<td>2-14</td>
<td>CATAGCTCAGTACGTCCAGCCATTTAGGAGT</td>
</tr>
<tr>
<td>2-15</td>
<td>GCCTTGCGTATAATATTGCCCATGGCTAGAGGAGTCGGATTG</td>
</tr>
<tr>
<td>2-16</td>
<td>ATGAATGCTCATCCGGAAATTCATAGTGAGGAGATCGATATGAGCGGGCCTCACG</td>
</tr>
<tr>
<td>2-17</td>
<td>GCCTTGCGTATAATATTGCCCATGGCTAGAGGAGTAAGATAGAGCTCTAAAAATCCAAGAAGTCAGCATTAGGAAAG</td>
</tr>
<tr>
<td>2-18</td>
<td>ATGAATGCTCATCCGGAAATTCATAGTGAGGAGATCGATATGAGCGGGCCTCACG</td>
</tr>
<tr>
<td>2-19</td>
<td>GCCTTGCGTATAATATTGCCCATGGCTAGAGGAGTAAGATAGAGCTCTAAAAATCCAAGAAGTCAGCATTAGGAAAG</td>
</tr>
<tr>
<td>2-20</td>
<td>ATGAATGCTCATCCGGAAATTCATAGTGAGGAGATCGATATGAGCGGGCCTCACG</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Sequence (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-21</td>
<td>GCCTTGCGTATAAATATTGCCCATGGTCAGTCATATAAATTTCCTAATTTTCTAAAATTTTCTTACCTGTCAGC</td>
</tr>
<tr>
<td>2-22</td>
<td>ATGAATGCTCATCCQGAATCTCTAGTGAAAGAATCGATATGAAAGATGTGTTAGATGATAACTTGGTAGAAAGACCGCT</td>
</tr>
<tr>
<td>2-23</td>
<td>ATCATTACTGCCTGAGACTGCACCGTGCAACCAA</td>
</tr>
<tr>
<td>2-24</td>
<td>GTCAACAAGCCCCGGATAGGGGTTCGCCGCACA</td>
</tr>
<tr>
<td>2-25</td>
<td>GAATACGCATCTCGAGGCTTGAACGTATTTAAGAAAG</td>
</tr>
<tr>
<td>2-26</td>
<td>CACGTATGCGGGATCCTAAATATACCTCTTTAATTACCTAGAAAGGAG</td>
</tr>
<tr>
<td>2-27</td>
<td>CACGTATGCGGGATCCATGAGTCACCCAGATCTA</td>
</tr>
<tr>
<td>2-28</td>
<td>GAATACGCTATCTCGAGTTAGTAATCTTTTGTCATG</td>
</tr>
<tr>
<td>2-29</td>
<td>CACGTATGCGGGATCCATGCGGTCATACGTCGCT</td>
</tr>
<tr>
<td>2-30</td>
<td>GAATACGCTATCTCGAGTCAATCACGACAGACGCC</td>
</tr>
<tr>
<td>2-31</td>
<td>CACGTATGCGGGATCCATGGAAGTAGAAAGAGTTATTATCTGATGAAGCAGAAACAAAGATTAC</td>
</tr>
<tr>
<td>2-32</td>
<td>GAATACGCTATCTCGAGCTATTTTGTTTCAACTTTATCTTTCCATTTCATTAAAAGAAGGCGT</td>
</tr>
<tr>
<td>2-33</td>
<td>CACGTATGCGGGATCCATGACTGAAATTTTTTCTGGTAGAAGGAGAGGTG</td>
</tr>
<tr>
<td>2-34</td>
<td>GAATACGCTATCTCGAGTTATAGAATCTAATTATTTGACGTAAACTTTAATAAAACATGATTACTTTGAGGATATG</td>
</tr>
<tr>
<td>2-35</td>
<td>GCAGAAGTTTATTTTGTT</td>
</tr>
<tr>
<td>2-36</td>
<td>GCCAACACAATACATC</td>
</tr>
</tbody>
</table>
8.2.4 Oligonucleotides used in Chapter 6

Table 8-4. Oligonucleotides used for cloning in chapter 6. For oligonucleotides with highlighted nucleotides: the nucleotides in bold black are the overlapping sequences with the parental vector; the nucleotides in italics are the restriction site; the nucleotides in bold blue are the RBS; the nucleotides in regular blue are random spacers; the nucleotides in bold orange are the restriction sites; the underlined nucleotides are the sequences homologous to the gene or sequence to amplify; and the regular letters are additional nucleotides added to allow restriction enzyme binding.

<table>
<thead>
<tr>
<th>Oligo #</th>
<th>Sequence (5' → 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-01</td>
<td>GAAATGTGCAGGAACCTATCTCGGCAGTCCTAGGCTCGGCTCAACGAGAACAAAC</td>
</tr>
<tr>
<td>3-02</td>
<td>GGTAGACACCGGGGTTCGCCCAGGGCTTCACGACACCAAC</td>
</tr>
<tr>
<td>3-03</td>
<td>GATATAACCCGGGTCCCTCCTCCTGCAAGAT</td>
</tr>
<tr>
<td>3-04</td>
<td>CTACATGATGTACAAATGATGAAGATCTCCAGG</td>
</tr>
<tr>
<td>3-05</td>
<td>GATCAGCCCGGATGACCATGACAATGCACCGTCA</td>
</tr>
<tr>
<td>3-06</td>
<td>GATCTAGATCTGTACATTGATGACCATGACAATGACAATGACAATGACCAATGACCATGACCAATGACCAATGACCAAT</td>
</tr>
<tr>
<td>3-07</td>
<td>GACGTTATCCCGGATGCGAGCTCTCAGGAT</td>
</tr>
<tr>
<td>3-08</td>
<td>GATCCAGTGTCATTATTACGTCGCTTTTCG</td>
</tr>
<tr>
<td>3-09</td>
<td>GACGTATCCCGGATGAGAAGCAACGTACGACCAAGCAACGTACGACCAAGCAACGTACGACCAAGCAACGTACGACCAAG</td>
</tr>
<tr>
<td>3-10</td>
<td>GATCCAGTGTCACATTAGCCAGCCGCGC</td>
</tr>
</tbody>
</table>
### 8.3 PCR conditions

Table 8-5. Conditions for every PCR performed in Chapter 3. All PCRs were performed with Q5 polymerase, except for colony PCRs, which were performed with Taq polymerase.

<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exonuclease III gene amplification to clone into pTTQ18 to obtain pQR1866</td>
<td>#1-01, #1-02</td>
<td>69 °C</td>
<td>1 min 27 s</td>
<td>E. coli MG1655 genomic DNA</td>
</tr>
<tr>
<td>λ exonuclease gene amplification to clone into pTTQ18 to obtain pQR1867</td>
<td>#1-03, #1-04</td>
<td>72 °C</td>
<td>20 s</td>
<td>λ bacteriophage genomic DNA</td>
</tr>
<tr>
<td>T5 exonuclease gene amplification to clone into pTTQ18 to obtain pQR1868</td>
<td>#1-05, #1-06</td>
<td>68 °C</td>
<td>26 s</td>
<td>T5 bacteriophage genomic DNA</td>
</tr>
<tr>
<td>Exonuclease III gene amplification to clone into pTTQ18 to obtain pQR1869</td>
<td>#1-01, #1-09</td>
<td>68 °C</td>
<td>27 s</td>
<td>E. coli MG1655 genomic DNA</td>
</tr>
<tr>
<td>Exonuclease III gene amplification to clone into pTTQ18 to obtain pQR1870</td>
<td>#1-01, #1-10</td>
<td>68 °C</td>
<td>36 s</td>
<td>E. coli MG1655 genomic DNA</td>
</tr>
<tr>
<td>λ exonuclease gene amplification to clone into pTTQ18 to obtain pQR1869</td>
<td>#1-04, #1-11</td>
<td>72 °C</td>
<td>22 s</td>
<td>λ bacteriophage genomic DNA</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>T5 exonuclease gene amplification to clone into pTTQ18 to obtain pQR1870</td>
<td>#1-06, #1-12</td>
<td>68 °C</td>
<td>28 s</td>
<td>T5 bacteriophage genomic DNA</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1866 colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 20 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pQR1866</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1867 colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 12 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pQR1867</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1868 colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 24 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pQR1868</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1869 colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>2 min 2 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pQR1869</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1870 colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>2 min 14 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pQR1870</td>
</tr>
</tbody>
</table>
Table 8-6. Conditions for every PCR performed in Chapter 5. All PCRs were performed with Q5 polymerase, except for colony PCRs, which were performed with Taq polymerase.

<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA, comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification of the <em>Pvu</em>II gene to clone into pQR445</td>
<td>#2-01, #2-02</td>
<td>67 °C</td>
<td>17 s</td>
<td><em>P. vulgaris</em> plasmid pPVU1</td>
</tr>
<tr>
<td>Amplification of the tetR cassette from pACYC184 to clone into pPVU1 to generate pPVU1</td>
<td>#2-03, #2-04</td>
<td>69 °C</td>
<td>40 s</td>
<td>pACYC184</td>
</tr>
<tr>
<td>NotI gene amplification to clone into pTTQ18</td>
<td>#2-05, #2-06</td>
<td>58 °C</td>
<td>16 s</td>
<td><em>N. otiditis-caviarum</em> genomic DNA</td>
</tr>
<tr>
<td><em>Pvu</em>II gene amplification to clone into pTTQ18</td>
<td>#2-07, #2-08</td>
<td>68 °C</td>
<td>29 s</td>
<td><em>P. vulgaris</em> plasmid pPVU1</td>
</tr>
<tr>
<td><em>BamHI</em> gene amplification to clone into pTTQ18</td>
<td>#2-09, #2-10</td>
<td>66 °C</td>
<td>42 s</td>
<td><em>B. amyloliquefaciens</em> genomic DNA</td>
</tr>
<tr>
<td><em>HpaII</em> gene amplification to clone into pTTQ18</td>
<td>#2-11, #2-12</td>
<td>69 °C</td>
<td>40 s</td>
<td><em>H. influenzae</em> genomic DNA</td>
</tr>
<tr>
<td>Amplification of pPVU1t to delete the <em>Pvu</em>II restriction enzyme gene and introduce a <em>Pvu</em>II site</td>
<td>#2-13, #2-14</td>
<td>59 °C</td>
<td>2 min 45 s</td>
<td>pPVU1t</td>
</tr>
<tr>
<td>Colony PCR to screen positive pTTQ18-NotI colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 36 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pTTQ18-NotI</td>
</tr>
<tr>
<td>Colony PCR to screen positive pTTQ18-<em>Pvu</em>II colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>56 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pTTQ18-*PvuII</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA, comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colony PCR to screen positive pTTQ18-BamHI colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 6 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pTTQ18-BamHI</td>
</tr>
<tr>
<td>Colony PCR to screen positive pTTQ18-HpaII colonies</td>
<td>#1-07, #1-08</td>
<td>54 °C</td>
<td>1 min 31 s</td>
<td>TOP10 colonies transformed with a CPEC reaction to clone pTTQ18-HpaII</td>
</tr>
<tr>
<td>NcoI methyltransferase gene amplification to clone into pACYC184</td>
<td>#2-15, #2-16</td>
<td>71 °C</td>
<td>1 min 8 s</td>
<td>Intermediate plasmid harbouring the methyltransferase (not mentioned in this work)</td>
</tr>
<tr>
<td>Pvull methyltransferase gene amplification to clone into pACYC184</td>
<td>#2-17, #2-18</td>
<td>68 °C</td>
<td>1 min 5 s</td>
<td>Intermediate plasmid harbouring the methyltransferase (not mentioned in this work)</td>
</tr>
<tr>
<td>BamHI methyltransferase gene amplification to clone into pACYC184</td>
<td>#2-19, #2-20</td>
<td>68 °C</td>
<td>1 min 20 s</td>
<td>Intermediate plasmid harbouring the methyltransferase (not mentioned in this work)</td>
</tr>
<tr>
<td>HpaII methyltransferase gene amplification to clone into pACYC184</td>
<td>#2-21, #2-22</td>
<td>68 °C</td>
<td>1 min 9 s</td>
<td>Intermediate plasmid harbouring the methyltransferase (not mentioned in this work)</td>
</tr>
<tr>
<td>Amplification of the exonuclease expression cassette from pQR1870 to clone into the recombinase controller plasmid</td>
<td>#2-23, #2-24</td>
<td>70 °C</td>
<td>2 min 10 s</td>
<td>pQR1870</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA, comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification of the expression plasmid to exclude the mkate2 gene and introduce <em>Bam</em>HI and <em>Xho</em>I sites to generate pQR1872, pQR1873, pQR1874 and pQR1875</td>
<td>#2-25, #2-26</td>
<td>58 °C</td>
<td>3 min</td>
<td>Expression plasmid (from the dual recombinase expression system). DMSO added to the PCR mix at 5%</td>
</tr>
<tr>
<td><em>Not</em>I gene amplification to clone into the expression plasmid</td>
<td>#2-29, #2-30</td>
<td>58 °C</td>
<td>16 s</td>
<td><em>N. otiditis-caviarum</em> genomic DNA</td>
</tr>
<tr>
<td><em>Bam</em>HI gene amplification to clone into the expression plasmid</td>
<td>#2-31, #2-32</td>
<td>68 °C</td>
<td>29 s</td>
<td><em>P. vulgaris</em> plasmid pPVU1</td>
</tr>
<tr>
<td><em>Pvu</em>II gene amplification to clone into the expression plasmid to generate pQR1872</td>
<td>#2-27, #2-28</td>
<td>66 °C</td>
<td>42 s</td>
<td><em>B. amyloliquefaciens</em> genomic DNA</td>
</tr>
<tr>
<td><em>Hpa</em>II gene amplification to clone into the expression plasmid to generate pQR1873</td>
<td>#2-33, #2-34</td>
<td>69 °C</td>
<td>40 s</td>
<td><em>H. influenzae</em> genomic DNA</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1872 colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>44 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to clone pQR1872</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA, comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colony PCR to screen positive pQR1873 colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>1 min 20 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to clone pQR1873</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1874 colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>34 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to clone pQR1874</td>
</tr>
<tr>
<td>Colony PCR to screen positive pQR1875 colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>41 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to clone pQR1875</td>
</tr>
<tr>
<td>Colony PCR to screen positive expression plasmid with NotI colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>1 min 24 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to obtain the expression plasmid harbouring NotI</td>
</tr>
<tr>
<td>Colony PCR to screen positive expression plasmid with BamHI colonies</td>
<td>#2-35, #2-36</td>
<td>56 °C</td>
<td>54 s</td>
<td>DH5αZ1 colonies transformed with a ligation reaction to obtain the expression plasmid harbouring BamHI</td>
</tr>
</tbody>
</table>
Table 8-7. Conditions for every PCR performed in Chapter 6. All PCRs were performed with Q5 polymerase.

<table>
<thead>
<tr>
<th>PCR reaction</th>
<th>Oligos</th>
<th>Annealing temperature</th>
<th>Extension time</th>
<th>Template DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification of the sfGFP constitutive expression cassette from pQR1876 to clone into the Smal site of pQR1871 in order to obtain pQR1877</td>
<td>#3-01, #3-02</td>
<td>67 °C</td>
<td>28 s</td>
<td>pQR1876</td>
</tr>
<tr>
<td>Amplification of the pQR1877 backbone to introduce Smal and BsrGI sites</td>
<td>#3-03, #3-04</td>
<td>65 °C</td>
<td>5 min 1 s</td>
<td>pQR1877</td>
</tr>
<tr>
<td>Amplification of the portion of the <em>Pseudomonas</em> meta pathway to introduce Smal and BsrGI sites in order to clone into pQR1877 to obtain pQR1878</td>
<td>#3-05, #3-06</td>
<td>66 °C</td>
<td>2 min 37 s</td>
<td>pQR226</td>
</tr>
<tr>
<td>Amplification of the tinsel purple gene to introduce Smal and BsrGI sites in order to clone into pQR1877 to obtain pQR1879</td>
<td>#3-07, #3-08</td>
<td>68 °C</td>
<td>22 s</td>
<td>Tinsel purple Km(^R) plasmid (ATUM)</td>
</tr>
<tr>
<td>Amplification of the <em>C. violaceum</em> CV2025 transaminase (TAm) to introduce Smal and BsrGI sites in order to clone into pQR1877 to obtain pQR1880</td>
<td>#3-09, #3-10</td>
<td>71 °C</td>
<td>42 s</td>
<td>pQR801</td>
</tr>
</tbody>
</table>
8.4 gBlock sequences

This section compiles the different sequences ordered as gBlocks (IDT, Leuven, Belgium) for cloning, together with a description of each part of the sequence.

8.4.1 ccdB

The nucleotides highlighted in orange are the BamHI and XhoI restriction sites, respectively; the underlined nucleotides are the ccdB CDS; the non-highlighted nucleotides are random nucleotides included to allow restriction enzyme recognition of the sites for cleavage.

CACGTATGCGGGATCCATGCGATTTAAGGTTTACACCTATAAAAGAGAGAGGCCGT
TATCGTTCTTGTGTGGATGTACAGAGTGATATTATTGACACGCCTGCCGGACGGGA
TGTTGATCCCCCTGCGCCAGTCAGCTGCTGCTGCTGCTACAGATTAAGTCTCCGGTGAACT
TTACCCGCTGGTGGATATCCGGGGATGAAAGCTGGGCATGATGACCCCGCCGGAT
GGCCAGTGTGCCCCTTGCTGCCATGTTGCGGGAAGAGTGGCTGTACAGCCGACGCGAG
CGAAAATGACATCAAAAAACGCCATTAACCTGATGTGTTCTGGGAATATAACTCGAGA
TGGTTATTC

8.4.2 denA

The nucleotides highlighted in orange are the BamHI and XhoI restriction sites, respectively; the underlined nucleotides are the denA CDS; the non-highlighted nucleotides are random nucleotides included to allow restriction enzyme recognition of the sites for cleavage.

CACGTATGCGGGATCCATGCGATTTAAGGTTTACACCTATAAAAGAGAGAGGCCGT
ACAGAATATTCTATTATATATATAGCTAGAAATTAGACACGCAATGGAAGTATA
AAAACATTATCTATTATATATATATATATATATATATATATATATATATATATATATAT
GAGCTTTTTATATGTGGAAAAACTAAAAAATTTACGTTAAAAGAAATACACTATTATAGA
ACTGCTATTACCGCAAGACACAAACTGCTGATTCTACTAAATCTGCATTATTCCAT
TCTGCCTAAAGGAGGAAGCAAAAGTTGAATTTTACGCCCGCCTAATGTTTAATCT
TTCTATGACAATAATGGAGTTAGGTACAATGACAATCGCAACGATTGACTTAGAGGGAC
CTCTATTTATTAACCTGGAAATATTCAACACACAGAAAAATGAC
TCGAGATGCGTATTC

8.4.3 sfGFP into pACYC184

The nucleotides in bold black are the overlapping sequences with pACYC184; the nucleotides highlighted in orange are the Ncol, Xhol and BglI restriction sites, respectively; the nucleotides in purple are the sequence for the J23119 promoter; the nucleotides in italics are random spacers between promoter and RBS; the nucleotides in green are the RBS; the underlined nucleotides are the sfGFP CDS.

GCACCTTTGTCGCTTGGCCTATAATATTGCACATGGTTGACAGCTAGCTCAGTCC
TAGGTATAATGCTAGCAGATACTTGCAGGAGGAGGACTCGAGATCGGTAAAGGCG
AAGAGCTGTTCACTGGGTGTCGTCCTATTCTGGTGAACGTGATTGATGTGATGCTAA
CGGTCATAAGTTTCCCCGTCGCTGGCCGAGGGTGGAAGGTGACGCACTTGGCCTGACTC
TGGTAACGACGCTGACTTTATGTTGCTTCGCTTTGCTCGTTATCCGGACCATATG
AAGCAGCATGACTTCTTCCAAGTCCGCCATTCGCCGGAAGGTATATGTGCAGGAAGCGCA
CGATTTCCTTAAAGGATGACGGCCAGCTACAAACCATGTCGGGAAGTAGAATTTGGA
AGGCAGTACCCTGGTAAACCCGCTTGGAGCTGAAAGGCATTGACTTTAAGAC
GGCAATATCCTGGGCCATAAGCTGGGAATACAAATTTTTACAGCCAAATTTTATCTG
CACGCCCGATAAAACAAAAATGCGATTAAACGCAATTTTTAAAAACCCGACAAACGC
TGGAGATGGCAGCGTCAGCTGGCTGATCACTACCAGCAAAACACTCCAATCG
GTGATGGTCCTTTCTGCTGACCAGCATCAGGATCTGCTATGATTGATGCTGGAAGGCT
GTCTAAAGATCCGAACGAGAAACCGCATATGGGATATGCTGCTGCTGCTGCTGCTGCT
GCAGCGGGGCATACGCGCATGGATGGTTGAGACACTTGCTCAGTTTTTAGAGATCTCCAG
GCATCAAATAACGGGAAAGGGCTCAGTGCAAGACGACTGGGCGCTTTTATCTG
TTGTTTGTGGTGAACGCTCTCAATTCGGATGAGCATTCATCAGGCGGGCAA
GAAT

8.4.4 Beginning of CV2025 TAm spanning the His-tag

The nucleotides in orange are the XhoI and MauBI sites respectively; the nucleotides in blue correspond to a glycine and six histidine residues; the rest is the sequence of pQR1880.

AGGAGGAGGACTCGAGATGGCCATCATCATCATCATCATCATATGCGAGGAAGCAACGT
ACGACCAGCCTGAGCCGCAACTGGATGCCGCCCATCACCTGCATCCGTTCACC
GATACCGCATCGCTGAACAGGCGGGCGCGCGCGTAATGACTAT
8.5 Exonuclease activity assay gels

Figure 8-29. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-30. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-31. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-32. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-33. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-34. Activity tests of protein extracts from E. coli CSR603 pTTQ18 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-35. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-36. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-37. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-38. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-39. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-40. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1866 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-41. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-42. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-43. Activity tests of protein extracts from induced E. coli CSR603 pQR1866 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-44. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-45. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-46. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-47. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-48. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-49. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1867 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-50. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-51. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-52. Activity tests of protein extracts from induced E. coli CSR603 pQR1867 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-53. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-54. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-55. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-56. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-57. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-58. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1868 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-59. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-60. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-61. Activity tests of protein extracts from induced E. coli CSR603 pQR1868 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-62. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-63. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-64. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-65. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-66. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-67. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1869 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-68. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-69. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-70. Activity tests of protein extracts from induced E. coli CSR603 pQR1869 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-71. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-72. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-73. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 1) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-74. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-75. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on BpmI-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-76. Activity tests of protein extracts from non-induced E. coli CSR603 pQR1870 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-77. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on EcoRV-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.

Figure 8-78. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on Bpml-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
Figure 8-79. Activity tests of protein extracts from induced E. coli CSR603 pQR1870 (replicate 2) on HindIII-digested λ DNA at different culture timepoints (3, 4, 5, 7, 9, 12, 24 h) and for different times (1, 5, 15, 30 min). The lane labelled as C is an equal amount of λ DNA with no protein extract.
8.6 Total DNA extraction gels

Figure 8-80. DNA gels with samples from different E. coli CSR603 cultures to which PvuII and exonuclease expression was induced and MMC and CI added, alongside the controls. n/a means no antibiotic addition; n/I means non-induced; Ind, induced; low, med and high allude to the different MMC, CI concentrations used; r.1 or r.2 are the different replicates for each condition.
Figure 8-81. DNA gels with samples from different E. coli DH5αZ1 cultures to which PvuII and exonuclease expression was induced and MMC and CI added, alongside the controls. n/I means non-induced; Ind, induced; low, med and high allude to the different MMC, CI concentrations used; r.1 or r.2 are the different replicates for each condition.
Table 8.8. DNA gels with samples from different E. coli DH5αZ1 pQR1871 pQR1872 cultures at different induction conditions, experiment discussed in Chapter 5, section 5.2.3. The Pvull, exo (3 h) sample has an additional timepoint because it was run alongside a different experiment.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Replicate #1</th>
<th>Replicate #2</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Pvull</em> (3 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Exo</em> (3 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Pvull</em> (7 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Exo</em> (7 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Pvull</em> (3 h), <em>exo</em> (7 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Exo</em> (3 h), <em>Pvull</em> (7 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Pvull</em>, <em>exo</em> (3 h)</td>
<td><img src="3h_5h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_5h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td><em>Pvull</em>, <em>exo</em> (7 h)</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
<tr>
<td>Non-induced</td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
<td><img src="3h_7h_11h_24h.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Table 8-9. DNA gels with samples from E. coli DH5αZ1 and CSR603 in Chapter 5, sections 5.2.3 to 5.2.5. The numbers above the lanes indicate the timepoint.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Replicate #1</th>
<th>Replicate #2</th>
<th>Replicate #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5αZ1 control</td>
<td><img src="image1" alt="" /></td>
<td><img src="image2" alt="" /></td>
<td><img src="image3" alt="" /></td>
</tr>
<tr>
<td>DH5αZ1</td>
<td><img src="image4" alt="" /></td>
<td><img src="image5" alt="" /></td>
<td><img src="image6" alt="" /></td>
</tr>
<tr>
<td>pQR1871</td>
<td><img src="image7" alt="" /></td>
<td><img src="image8" alt="" /></td>
<td><img src="image9" alt="" /></td>
</tr>
<tr>
<td>pQR1872</td>
<td><img src="image10" alt="" /></td>
<td><img src="image11" alt="" /></td>
<td><img src="image12" alt="" /></td>
</tr>
<tr>
<td>DH5αZ1</td>
<td><img src="image13" alt="" /></td>
<td><img src="image14" alt="" /></td>
<td><img src="image15" alt="" /></td>
</tr>
<tr>
<td>pQR1871</td>
<td><img src="image16" alt="" /></td>
<td><img src="image17" alt="" /></td>
<td><img src="image18" alt="" /></td>
</tr>
<tr>
<td>pQR1873</td>
<td><img src="image19" alt="" /></td>
<td><img src="image20" alt="" /></td>
<td><img src="image21" alt="" /></td>
</tr>
<tr>
<td>DH5αZ1</td>
<td><img src="image22" alt="" /></td>
<td><img src="image23" alt="" /></td>
<td><img src="image24" alt="" /></td>
</tr>
<tr>
<td>pQR1871</td>
<td><img src="image25" alt="" /></td>
<td><img src="image26" alt="" /></td>
<td><img src="image27" alt="" /></td>
</tr>
<tr>
<td>pQR1874</td>
<td><img src="image28" alt="" /></td>
<td><img src="image29" alt="" /></td>
<td><img src="image30" alt="" /></td>
</tr>
<tr>
<td>DH5αZ1</td>
<td><img src="image31" alt="" /></td>
<td><img src="image32" alt="" /></td>
<td><img src="image33" alt="" /></td>
</tr>
<tr>
<td>pQR1871</td>
<td><img src="image34" alt="" /></td>
<td><img src="image35" alt="" /></td>
<td><img src="image36" alt="" /></td>
</tr>
<tr>
<td>pQR1875</td>
<td><img src="image37" alt="" /></td>
<td><img src="image38" alt="" /></td>
<td><img src="image39" alt="" /></td>
</tr>
<tr>
<td>CSR603 control</td>
<td><img src="image40" alt="" /></td>
<td><img src="image41" alt="" /></td>
<td><img src="image42" alt="" /></td>
</tr>
<tr>
<td>CSR603</td>
<td><img src="image43" alt="" /></td>
<td><img src="image44" alt="" /></td>
<td><img src="image45" alt="" /></td>
</tr>
<tr>
<td>pQR1871</td>
<td><img src="image46" alt="" /></td>
<td><img src="image47" alt="" /></td>
<td><img src="image48" alt="" /></td>
</tr>
<tr>
<td>pQR1872</td>
<td><img src="image49" alt="" /></td>
<td><img src="image50" alt="" /></td>
<td><img src="image51" alt="" /></td>
</tr>
</tbody>
</table>
Bibliography


available at: https://jemi.microbiology.ubc.ca/sites/default/files/Bennett et al..pdf.


Cases, I. and de Lorenzo, V. (2005), “Genetically modified organisms for the environment: Stories of success and failure and what we have learned from them”, International Microbiology, 8, pp. 213–222.


315


Ritger, K., Black, S., Weaver, K., Jones, J., Gerber, S., Conover, C., Soyemi, K., et al.


