CHARACTERISATION OF AN EXPRESSION SYSTEM FOR COMMERCIAL PRODUCTION OF PROTEINS

by

LENE JØRGENSEN

Thesis submitted for the degree of Doctor of Philosophy

in

The University of Adelaide
Faculty of Engineering
and
Faculty of Science

May 1997

SUMMARY

To enable the production of recombinant chloramphenicol acetyl transferase (CAT) in *Escherichia coli* JM101, a number of *promoter*::*cat* transcriptional fusions were constructed. These fusions contained either IPTG-inducible or stationary-phase inducible promoters in a bidirectional promoter probe vector. Recombinant protein expression in the engineered systems was quantified using validated assays developed in this thesis. Furthermore, mathematical models were employed to establish the relative efficiencies of transcription and translation.

The detection of bacterially-expressed *cat* mRNA by slot-blotting was found to be highly dependent on total RNA immobilised onto the solid support, as well as mRNA concentration. mRNA quantitation by comparison with a pure standard resulted in gross underestimation because of this possible steric hindrance. A new method to quantitate *cat* mRNA was therefore developed. The new protocol for *cat* mRNA detection included a three-dimensional standard calibration curve, constructed for each assay, and overcomes the confounding effect of contaminating RNA.

The French press was more efficient at disrupting cells and releasing proteins than sonication. French pressing disrupted all cells in suspension whereas a maximum of 80% of the cells were disrupted following sonication. The level of CAT release was highest when cells were totally disrupted. Additional treatment with the detergent Triton X-100 was necessary to maximise CAT recovery.

Promoters induced by IPTG are commonly used but have both cost and environmental penalties. Nevertheless, an IPTG-induced system was included in this work as a control, to compare the relative efficiency of a system with commercial potential (stationary phase promoter system). Maximal protein expression was achieved for 0.1 mM IPTG after induction at OD_{600} = 0.8 in both shake-flask and fermentation experiments. A concentration of 0.4 mM IPTG yielded maximal expression for induction at OD_{600} = 2.4. Maximum CAT protein expression was independent of oxygen concentration. However, CAT protein production was highly dependent on the growth phase of the culture at induction. Induction close to stationary phase produced lower levels of CAT compared to

induction in logarithmic phase. Also, inoculation with a stationary-phase culture gave better CAT protein yield than fermenters inoculated with a logarithmic phase culture.

CAT protein production under control of the *tac* promoter was clearly limited at the translational level. This was shown by constant CAT protein levels after induction for decreasing ribosomal (16S rRNA) levels. Furthermore, induction with IPTG concentration beyond optimal resulted in a concomitant increase in mRNA level but not CAT protein. Translational limitation was confirmed by a simple mathematical model to establish the relative efficiencies of transcription (16-81 %) and translation (0.2-1 %).

Separation of the growth and CAT production phase was achieved using the stationary phase *katE* gene promoter. Batch fermentation experiments in minimal media showed that the transition from logarithmic growth phase to stationary phase stimulated *katE* expression. However, despite published material to the contrary, no inducers of the *katE* gene promoter were identified in minimal media. Acetate and o-hydroxybenzoate did not stimulate promoter activity. O-hydroxybenzoate actually inhibited translational activities. Glucose addition did improve CAT protein levels, but this was probably due to increased translational activity.

Final CAT protein levels per cell for the IPTG inducible *tac* promoter system were about 250 times higher than the stationary-phase inducible *katE* gene promoter in batch fermentation experiments. However, CAT protein levels under control of the *katE* gene promoter was limited at the translational level in batch experiments. Protein yield was improved by continuously feeding glucose, thus establishing a growth-limited culture for CAT protein production. A 3-fold increase in CAT protein, and a 10-fold increase in *cat* mRNA levels, was achieved. Future work to define a critical growth rate below which the stringent control is induced, may allow the *katE* gene promoter to be employed for commercial purposes. Furthermore, more information regarding promoter regulation is needed.

TABLE OF CONTENTS

STATEMENT OF ORIGINALITY i
ACKNOWLEDGEMENTS ii
SUMMARY iii
PUBLICATIONS ARISING FROM THIS THESIS RESEARCH v
ABBREVIATIONS vi

CHAPTER 1. INTRODUCTION

1.1 Recombinant DNA technology 2
1.2 Hosts for recombinant expression systems 3
1.3 *Escherichia coli* expression systems 4
1.3.1 Recombinant protein production under control of stationary phase inducible promoters 8
1.4 Reporter gene 10
1.5 Maximising production of recombinant protein 11
1.6 Specific aims of this thesis 14

CHAPTER 2. MATERIALS AND METHODS

2.1 Chemicals and reagents 19
2.2 Growth media 19
2.3 *E. coli* bacterial strains used in this study 21
2.4 Plasmids used in this study 21
2.5 Oligonucleotides 24
2.6 Enzymes. 24
2.7 Transformation 24
2.8 DNA extraction procedures 26
2.9 Analysis and manipulation of DNA 27
2.9.1 DNA quantitation 27
2.9.2 Agarose gel electrophoresis of DNA 27
2.9.3 Isolation and purification of gene fragments 28
2.9.4 Calculation of restriction fragment size 28
2.9.5 Dephosphorylation of DNA using alkaline phosphatase 28
2.9.6 End-filling with Klenow fragment 29
2.9.7 *In vitro* cloning 29
2.10 Synthesis of oligodeoxynucleotides 29
2.11 Polymerase Chain Reaction (PCR) 29
2.12 DNA sequence analysis 30
2.13 RNA analysis 32
2.13.1 RNA quantitation 32
2.13.2 RNA extraction and purification 32
2.13.3 Slot-Blot hybridization analysis 33
2.13.4 Hybridisation 33
2.13.5 RNA detection by chemiluminescence 34
2.14 Synthesis of the sense strand of *cat* mRNA 34

2.15 Synthesis of DIG-labelled *cat* RNA probe 36
2.16 Synthesis of the sense strand of 16S rRNA 36
2.17 Synthesis of DIG-labelled 16S rRNA probe 36
2.18 Preparation of Dynabeads 38
2.19 Cell counts 38
2.20 Cell dry weight 38
2.21 Cell density 39
2.22 Cell disruption 39
2.23 Total CAT Protein 39
2.24 Total soluble proteins 41
2.25 Plasmid copy number 41
2.26 Fermentations 42
2.27 Glucose analysis 43
2.28 Acetate analysis 43

CHAPTER 3. PREPARATION AND NUCLEOTIDE SEQUENCE OF CONSTRUCTS.

3.1 Strategy for construction of reporter plasmids with *cat* 44
3.2 Construction of *katE*::*cat* transcriptional fusion (pCT100) 44
3.3 Construction of *katF*::*cat* transcriptional fusion (pCT101) 48
3.4 Construction of *trc*::*cat* transcriptional fusion (pCT102) 51
3.5 Construction of *tac*::*cat* transcriptional fusion (pCT103) 51
3.6 Construction of plasmids for *in vitro* production of DIG-labelled *cat* RNA probe and positive control 55
3.6.1 Construction of pCT121 55
3.7 Construction of plasmid for *in vitro* production of DIG-labelled 16S rRNA probe and positive control 60
3.7.1 Construction of pCT123 60

CHAPTER 4. OPTIMISATION AND VALIDATION OF ASSAYS

4.1 Introduction 67
4.2 Results and discussion 67
4.2.1 Recovery of CAT protein from *E. coli* 67
4.2.2 Recovery of total soluble protein from *E. coli* 75
4.2.3 CAT and total soluble protein integrity during sample storage 78
4.2.4 Total RNA isolation 81
4.2.5 RNA integrity after storage 82
4.2.6 Quantitation of *cat* mRNA extraction 86
4.2.7 *cat* mRNA integrity during extraction 92
4.2.8 Determination of halflife ($t_{1/2}$) for *cat* mRNA 93
4.2.9 Plasmid copy number 95
4.2.10 Cell count 97
4.3 Conclusion 100

CHAPTER 5. RECOMBINANT PROTEIN PRODUCTION UNDER CONTROL OF IPTG INDUCED PROMOTERS.

5.1 Introduction 103
5.2 Results and discussion 107
5.2.1 Shake flask experiments 107
5.2.2 Fermentation experiments 112
5.3 Model description 118
5.3.1 Estimation of model parameters 120
5.3.2 Model regressions 120
5.4 Conclusion 125

CHAPTER 6. RECOMBINANT PROTEIN PRODUCTION UNDER CONTROL OF STATIONARY PHASE INDUCED PROMOTERS.

6.1 Introduction 128
6.2 Results and discussion 131
6.3 Model description 148
6.3.1 Estimation of model parameters 149
6.3.2 Model regressions 149
6.4 Comparison of the *katE* gene or *tac* promoters 152
6.5 Conclusion 154

CHAPTER 7. SUMMARY AND CONCLUSION 156

Future work 163

APPENDIX A.1 165
NOMENCLATURE 166
BIBLIOGRAPHY 167