THERMAL INDUCTION AND OVEREXPRESSION OF A FOREIGN PROTEIN
IN ESCHERICHIA COLI CULTURE

by

FIRDAUSI RAZALI

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Chemical Engineering

Waterloo, Ontario, Canada, 2002

©Firdausi Razali 2002
I hereby declare that I am the sole author of this thesis.

Jidavisi Razali

I authorize the University of Waterloo to lend the thesis to other institutions or individuals for the purpose of scholarly research.

Jidavisi Razali

I further authorize the University of Waterloo to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Jidavisi Razali
The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.
ABSTRACT

Engineering studies were performed on foreign protein production by a recombinant *Escherichia coli*, which overexpresses intracellular *E.coli* β-galactosidase from the plasmid vector pBRG 401 under transcriptional regulation of the λp<sub>1</sub> promoter and cl<sub>857</sub> protein repressor. Both genetic/physiological and environmental/bioprocess factors that influenced the performance and the kinetics of protein expression and growth in complex medium were investigated. A novel approach in describing protein expression in the thermal induction system in an *E.coli* culture was proposed.

The disappearance of protein expression during induction was primarily due to plasmid instability. The instability was non-segregational, irreversible, and strongly temperature-dependent. The loss of resistance towards the selective pressure was in parallel with the decline of protein expression, despite the fact that both events were regulated by separate promoters. The absence of inclusion bodies implied that the overexpressed proteins were completely soluble. The occurrence of intracellular proteolysis was not evident during the induction. However, proteolysis was observed even in the presence of a respiratory inhibitor at extended incubation time if glucose was present. This proved that the additional metabolic energy needed for the protein degradation was minimal.

Although oxygen concentration was not a limiting factor in the culture, the increased secretion of acetic acid following protein overexpression signified a shift of metabolism from oxidative to fermentative pathway. This resulted in a feedback repression in the Kreb’s cycle, with consequently lowering the metabolic overload due to reduced protein expression on the plasmid. The accumulated acetic acid and the acidic pH lowered both growth and protein expression. However, in the absence of acetic acid, the growth seemed unaffected by pH. The exclusion of glucose during the induction phase in complex medium was beneficial. The average specific enzyme activity decreased from 95.3 U/mg in the absence of glucose to 20.5 U/mg when glucose was present at an induction temperature of 42 °C. Complex-substrate concentrations also
affected plasmid stability, and hence determined protein expression levels. One may conclude that the inactivation of plasmid was not due to temperature *per se*, but rather due to the negative effect of a strong promoter and protein expression on the plasmid.

The average specific productivity decreased as the induction time elapsed. The plateau in productivity marked the complete inactivation of plasmid, and hence the end of protein expression. The decline of the total viable cell counts during the induction was mainly due to the metabolic load. When the culture was totally freed from the productive population, the total viable cell count started to incline. The increase of biomass from the unproductive cells resulted in the decline of the total specific protein activity in the culture. The time of the incline coincided with the time where all the resistant (productive) cells completely disappeared from the culture.

A novel semi-empirical model was developed to describe the kinetics of protein expression, growth, and plasmid stability. Good agreement between the predicted and experimental results, in both instantaneous and ramped induction studies, demonstrated the applicability of the model. A novel approach to predicting the induction temperature operating range was established and experimentally proven between 37.3 °C and 42 °C in this study. The ultimate specific β-galactosidase activity ranged from 66 U/mg to 149 U/mg at induction temperature of 42 °C and 38 °C respectively. This approach also provided an accurate means of estimating the average plasmid number per cell at a given time and temperature during the induction. The dependence of parameters on the temperature was determined using an Arrhenius relationship. The temperature strongly affected the plasmid decay constant (λ) and the number of activated plasmids per cell (n), while the specific growth rate (μ) was relatively independent. This suggested that the protein expression and growth were uncoupled. In addition, the maximum biomass concentration (X_max), the proteolysis rate (k_prot), and lysis rate (k_lys) were also barely influenced by temperature.
Acknowledgements

I would like to express my deep gratitude to my supervisors, Professor M. Moo-Young, and Professor J.M. Scharer for their encouragement, guidance, and advice. It was a privilege to be able to work with them, as they are high calibre researchers. I would also like to record my thanks to Prof. Bernie Glick on his generous contributions to this study.

I am truly indebted to all my colleagues in both Biochemical Engineering Group and Prof. Glick’s Lab. members, particularly to Zairossani Nor, Ahmad Ramli Yahya, Edmund Kwong, Oscar Kibazohi, Melih Tamer, Otunola Adedayo, Sun-il Yun, Wanwisa Skolpap, Hong Yang, David Stuivenberg, Jiping, and Nikos Hontzeas for their advice and contributions to this work. I would like to acknowledge the kind assistance of Jana Otruba, Ralph Dickhout, Dale Weber (Biology) and Frank Wassmer in the analytical part and experimental setup of this research.

The financial support of Universiti Teknologi Malaysia (UTM) for this study is greatly appreciated. Other financial support from National Sciences and Engineering Research Council of Canada (NSERC) in a form of research assistantships and conference-related expenses is also acknowledged.

I also like to express my deepest gratitude to my beloved wife, Fareha Abdul Rahman, for her continuous encouragement, and patience in dealing with hardship throughout our living in Canada. Also, I wish to acknowledge my family members for their constant moral support throughout this study. I would like to dedicate this thesis to my beloved parents for being the source of inspiration. May Allah shower His mercy on them.

And last but not least, I would like to thank the readers of this thesis, Professor R.J. Neufeld, Professor W.A. Anderson, and Prof. E.J. Jervis for their constructive comments and contributions to this thesis.
# TABLE OF CONTENTS

ABSTRACT ........................................................................................................iv  
Acknowledgements ........................................................................................vi  
TABLE of CONTENTS ..................................................................................vii  
List of Tables ..................................................................................................xi  
List of Figures ..................................................................................................xii  
List of Appendices ..........................................................................................xvii  
NOMENCLATURE ..........................................................................................xx  

## CHAPTER 1: INTRODUCTION .................................................................1  
1.1 *Escherichia coli* as an Expression Host ................................................2  
1.2 Studies of Heterologous Protein Expression Using Regulatable Strong Promoters in *Escherichia coli* Cultivations ............................3  
1.3 Problems Associated with *p*\(^1\) Promoter System and Protein Overexpression in *Escherichia coli* ......................................................5  
1.4 Objectives ..............................................................................................9  

## CHAPTER 2: LITERATURE REVIEW ......................................................11  
2.1 Inducible System and Protein Overexpression under Transcriptional Regulation of *p*\(^1\) Promoter System ..................................................12  
2.2 Implications of Protein Overexpression in *Escherichia coli* Culture ......14  
  2.2.1 Inclusion Body Formation ...............................................................14  
  2.2.2 Metabolic Overload ........................................................................16  
  2.2.3 Stress-related Intracellular Proteolysis .........................................19  
  2.2.4 Increase of Acetic Acid Secretion ..................................................20  
2.3 Bioprocess Approaches in Improving Protein Expression in *p*\(^1\) Promoter System ..................................................................................26  
  2.3.1 Temperature Upshift Schemes .......................................................26  
  2.3.2 Feeding of Substrate .......................................................................28  
  2.3.3 Timing of Induction .......................................................................31  
  2.3.4 Two-stage Culture ..........................................................................32
CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION .............................................59
  4.1 Introduction .................................................................................................59
  4.2 Characteristics of Plasmid Stability ............................................................61
  4.3 Solubility and Bioactivity of the Expressed Protein ....................................65
  4.4 Stability of the Expressed Protein towards Intracellular Proteolytic Activity During and Post-induction .................................................................73
  4.5 Effect of Active Plasmid Presence and Temperature on Cell Viability and Specific Growth Rate ..........................................................77
  4.6 Oxygen Limitation during Induction ............................................................83
  4.7 Crabtree-like Effect of Carbon Sources on Growth and Protein Expression .................................................................87
  4.8 Inhibitory Effect of Acetic Acid and Its Interaction with pH on Growth and Protein Expression .................................................................94
  4.9 Effect of Substrate Concentrations and Induction Temperatures on Plasmid Stability and Protein Expression .............................................102
  4.10 Effect of Induction Temperature on Productivity, Growth and Plasmid Stability ..........................................................109

CHAPTER 5: MODELING OF ESCHERICHIA COLI ..............................................125
  5.1 Introduction .................................................................................................125
  5.2 Growth and Protein Expression .................................................................126
  5.3 Plasmid Inactivation ...................................................................................131
  5.4 Estimation Procedure ................................................................................134
  5.5 Analyses of Parameter Estimates ..............................................................140
  5.6 Comparison between Model Simulation and Experimental Results ........144
List of Tables

Table 1.1: Studies Employing $\lambda p^{l}/cl_{857}$ Promoter System in
Escherichia coli Cultures ................................................................. 6

Table 1.2: Bioprocess Strategies in Minimizing Plasmid Segregational
Instability ......................................................................................... 9

Table 3.1: Combination Table of Acrylamide Solution, Separating Gel
Buffer, and Deionized Water Volumes to form a Separating Gel ...... 54

Table 5.1: Values of Parameters at 32 °C, 38 °C, 40 °C, and 42 °C
Estimated from Model ................................................................. 140
List of Figures

Figure 2.1: Transmission Electronmicrograph of Alpha-consensus Human Interferon (IFN-α-Con) in Form of Inclusion Body within Escherichia coli Cells (Fieschko et al., 1985) .......................15

Figure 2.2: Glucose Metabolic Pathways of Escherichia coli .......................21

Figure 3.1: Schematic Diagram of Plasmid pBRG 401 ..................................43

Figure 4.1: Plasmid Instability in Culture Induced at 42 °C ..........................62

Figure 4.2: Determination of Post-induction (42 °C for 3 hours) Plasmid DNA Presence in Escherichia coli Cells by Agarose Gel Electrophoresis........................................64

Figure 4.3: Determination of Post-Induction Plasmid Recovery of 20 °C-incubated Culture by Colony Formation on Kanamycin-supplemented Agar (Shaded = Induced at 42 °C for 3 Hours; Unshaded = Uninduced) ........................................66

Figure 4.4: Comparison of Electronmicrographs of E. coli Cells before (left) and after (right) Induction (42 °C for 3 hours) for Determining the Presence of Inclusion Bodies ........................................67

Figure 4.5: Comparison of Cells between before (left) and after (right) Disruption at Maximum Volumetric Activity Using Glass Beads for Determining the Size of Cell Debris ........................................69

Figure 4.6: Determination of Released Protein from Disrupted Cells by Comparing Specific β-galactosidase Activity (U/mg) in Lysate, Supernatant of Lysate, and Shaken Lysate ........................................70

Figure 4.7: Time-course Images of β-galactosidase Bands on SDS Electrophoresis Gels from Cultures Induced at 38 °C and 42 °C in the Absence of Glucose (Top), and With and Without Glucose Induced at 42 °C (Bottom) ........................................72

Figure 4.8: Linear Regressions of Volumetric β-galactosidase Activities
to Trace Values (Intensity time path length in mm) Obtained
from Band Images on SDS Gels of Figure 4.7 ..................................74

Figure 4.9: Determination of Intracellular Proteolysis and its Interaction
with Endogenous Energy on β-galactosidase at 42°0C (20 mM
Sodium Azide = Respiratory Chain Inhibitor, 2.5 g/l Glucose
= Energy Source) ...............................................................76

Figure 4.10: Determination of Intracellular Proteolysis of β-galactosidase
and its Interaction with Endogenous Energy in 20°0C-incubated
Culture at 20-Hour Incubation ............................................77

Figure 4.11: Comparison of Viability (CFU/ml) of Kanamycin-sensitive
Cultures between Induction Temperature of 38°0C and 42°0C ......79

Figure 4.12: Comparison of Viability (CFU/ml) of Kanamycin-resistant
Cultures between Induction Temperature of 38°0C and 42°0C ......80

Figure 4.13: Trend of Average Specific Growth Rate (hour⁻¹) of
Kanamycin-resistant (R) and Kanamycin-sensitive (S) at
Temperature of 38°0C and 42°0C ...........................................82

Figure 4.14: Increase of Oxygen Consumption during Protein Overexpression
(at 42°0C) in 15-l MBR Aerated Bioreactor (Induction Begin =
Hour 9, Induction End = Hour 12) ........................................84

Figure 4.15: Comparison of Specific β-galactosidase Expression (U/mg)
between Cultures in 250-ml Erlenmeyer Flask (50 ml Broth) and
15-l MBR Aerated Bioreactor (5.0 l Broth) Induced at 42°0C ......86

Figure 4.16: Crabtree-like Effect on Specific β-galactosidase Expression
of Cultures Supplemented with 5.0 g/l Glucose and 5.0 g/l
Glycerol Induced at 42°0C ...................................................88

Figure 4.17: Enhanced Release of Acetic Acid due to Limitation of
Oxidative Pathway during Induction at 42°0C in Cultures
Supplemented with 5.0 g/l Glucose and 5.0 g/l Glycerol ..........90

Figure 4.18: Crabtree-like Effect on Final Percentage of Kanamycin-resistant
Cells in Induced (shaded) and Uninduced (unshaded) Cultures
Supplemented with 5.0 g/l Glucose and 5.0 g/l Glycerol ..........91
Figure 4.19: Crabtree-like Effect on Final Total Viable Cell Count (CFU per ml) in Induced (shaded) and Uninduced (unshaded) Cultures Supplemented with 5.0 g/l Glucose and 5.0 g/l Glycerol ......................... 92

Figure 4.20: Time-Course Profile of pH Changes due to Acetic Acid Release during Glucose and Glycerol Metabolisms following Induction at 42 °C ......................................................... 95

Figure 4.21: Comparison of Glucose Consumption Profile between Induced (at 42 °C) and Uninduced Cultures ............................................. 97

Figure 4.22: Inhibitory Effect of Acetic Acid and pH on Dry Biomass Concentration (g/l) during Induction at 42 °C ........................................ 99

Figure 4.23: Inhibitory Effect of Acetic Acid and pH on Volumetric β-galactosidase Expression (U/ml) during Induction at 42 °C .......... 100

Figure 4.24: Inhibitory Effect of Acetic Acid and pH on Specific β-galactosidase Expression (U/mg) during Induction at 42 °C .......... 101

Figure 4.25: Effect of Medium (LB) Concentrations on Percentage of Kanamycin-resistant Cells Induced at 42 °C ................................. 104

Figure 4.26: Stabilization of Plasmid by Substrate Feeding, Lower Induction Temperature (38 °C), and Ramped-temperature Upshift (6 °C/hour) ................................................................. 105

Figure 4.27: Comparison of Specific β-galactosidase Expression (U/mg) between Batch and Fed-batch Cultures Induced at 42 °C ............. 107

Figure 4.28: Comparison of Specific β-galactosidase Expression (U/mg) between Batch and Fed-batch Cultures Instantaneously Induced at 38 °C ................................................................. 108

Figure 4.29: Comparison of Specific β-galactosidase Expression (U/mg) between Instantaneous (42 °C) and Ramped-temperature Upshift (6 °C/hour) in Batch Cultures ......................................................... 110

Figure 4.30: Volumetric β-galactosidase Expression (U/ml) of Cultures Induced at 38 °C, 40 °C, and 42 °C ......................................................... 112

Figure 4.31: Percentage of Kanamycin-resistant Cells Induced at 38 °C, 40 °C, and 42 °C ................................................................. 113
Figure 4.32: Total Viable Cell Count (CFU/ml) of Cultures Induced at 32 °C, 38 °C, 40 °C, and 42 °C .................................................................114

Figure 4.33: Dry Biomass Concentration (g/l) of Cultures Induced at 38 °C, 40 °C, and 42 °C .................................................................115

Figure 4.34: Specific β-galactosidase Expression (U/mg) of Cultures Induced at 38 °C, 40 °C, and 42 °C .................................................................116

Figure 4.35: Influence of Temperature Apparent Specific Growth Rate (hour⁻¹) ....119

Figure 4.36: Effect of Induction Temperature (38 °C, 40 °C, and 42 °C) on Average Specific β-galactosidase Productivity (Umg⁻¹hour⁻¹) ..........120

Figure 4.37: Correlation between Initial Average Specific β-galactosidase Productivity (U mg⁻¹hour⁻¹) and Induction Temperatures .........121

Figure 4.38: Specific β-galactosidase Activity (U/mg) from Ramped-temperature Upshift (6 °C/hour) Experiment ..................................123

Figure 5.1: Dependence of Estimated Parameters on the Induction Temperatures ............................................................................142

Figure 5.2: Dependence of Average Initial Number of Active Plasmids per Cell (n) on Induction Temperatures ...............................145

Figure 5.3: Model Simulation of Time-Course Decay of Average Number of Active Plasmids per Cell (n) at Induction Temperatures of 38 °C, 40 °C, and 42 °C .................................................................146

Figure 5.4: Comparison of Volumetric β-galactosidase Activity (U/ml) at 38 °C, 40 °C, and 42 °C between Model Simulation (dashed-lines) and Experimental Data (markers) .................................................................150

Figure 5.5: Comparison of Percentage of Kanamycin-resistant Cells at 38 °C, 40 °C, and 42 °C between Model Simulation (dashed-lines) and Experimental Data (markers) .................................................................151

Figure 5.6: Comparison of Dry Biomass Concentration (g/l) of Cultures Induced at 38 °C, 40 °C, and 42 °C between Model Simulation (dashed-lines) and Experimental Data (markers) .................................................................152

Figure 5.7: Comparison of Specific β-galactosidase Activity (U/mg) at
38 °C, 40 °C, and 42 °C between Model Simulation (dashed-lines) and Experimental Data (markers) .........................................................153

Figure 5.8: Comparison of Dry Biomass Concentration (g/l) of Ramped-temperature Upshift (6 °C/hour) Experiment between Model Simulation and Experimental Data ........................................157

Figure 5.9: Comparison of Volumetric β-galactosidase Activity (U/ml) in Ramped-temperature Upshift (6 °C/hour) Experiment between Model Simulation and Experimental Data ........................................158

Figure 5.10: Comparison of Specific β-galactosidase Activity (U/mg) in Ramped-temperature Upshift (6 °C/hour) Experiment between Model Simulation and Experimental Data ........................................159
# List of Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>APPENDIX A</td>
<td>Summary of E.coli's β-galactosidase Characteristics</td>
<td>167</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>Gravimetric Analysis of Biomass</td>
<td>168</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>Temperature Profile in the Culture Broth during Instantaneous Induction from 30 °C to 42 °C</td>
<td>169</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>Preliminary Results</td>
<td></td>
</tr>
<tr>
<td>Appendix D1</td>
<td>Effect of pH, Medium Type, and Temperature on Protein Expression and Growth</td>
<td>170</td>
</tr>
<tr>
<td>Appendix D2</td>
<td>Medium M9 Composition</td>
<td>173</td>
</tr>
<tr>
<td>Appendix D3</td>
<td>Improvement of M9 medium by Strengthening Nitrogen Source from Cassamino Acid and Ammonium Chloride</td>
<td>174</td>
</tr>
<tr>
<td>Appendix D4</td>
<td>Effect of Cassamino Acid and Vitamins on Growth in Defined Medium M9</td>
<td>175</td>
</tr>
<tr>
<td>Appendix D5</td>
<td>Effect of Amino Acids on Growth in Defined Medium M9</td>
<td>176</td>
</tr>
<tr>
<td>Appendix D6</td>
<td>Effect of L-isoleucine, L-cysteine, L-tryptophan, L-valine on Growth</td>
<td>177</td>
</tr>
<tr>
<td>Appendix D7</td>
<td>Data of Stepwise Increasing Rate Fed-batch: Late Induction</td>
<td>178</td>
</tr>
<tr>
<td>Appendix D8</td>
<td>Data of Stepwise Increasing Rate Fed-batch: Early Induction</td>
<td>180</td>
</tr>
<tr>
<td>Appendix D9</td>
<td>Comparison of Data between Early and Late Induction of Fed-batch Cultures</td>
<td>182</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>Experimental Results</td>
<td></td>
</tr>
<tr>
<td>Appendix E1</td>
<td>Trace Values (Intensity time path length in mm) obtained from Band Images on SDS Gel Electrophoresis (Figure 4.7 and 4.8)</td>
<td>183</td>
</tr>
<tr>
<td>Appendix E2</td>
<td>Effect of Sodium Azide and Glucose on β-galactosidase Activity (Figure 4.9, Figure 4.10)</td>
<td>184</td>
</tr>
<tr>
<td>Appendix E3</td>
<td>Viable Cell Count of Kanamycin-resistant and Kanamycin-sensitive Cultures Grown at 30 °C, 38 °C and 42 °C (Figure 4.11, 4.12)</td>
<td>185</td>
</tr>
<tr>
<td>Appendix E4</td>
<td>Dry Biomass Concentration (g/l) of Kanamycin-resistant and Kanamycin-sensitive Cultures Grown at 30 °C, 38 °C</td>
<td></td>
</tr>
</tbody>
</table>
Appendix E5: Dry Biomass Concentration (g/l) and Percentage of Air Saturation in 15-l MBR Bioreactor (Figure 4.14) ..................187

Appendix E6: Specific β-galactosidase Activity in 250-ml Erlenmeyer Flask (50-ml broth) and 15-l MBR Bioreactor (6-l broth) (Figure 4.15) ..............................................................188

Appendix E7: Effect of 5.0 g/l Glycerol and 5.0 g/lGlucose on Specific β-galactosidase Activity (Figure 4.16) .................................189

Appendix E8: Release of Acetic Acid from Glycerol and Glucose Metabolisms in Induced (42 °C) and Uninduced Cultures (Figure 4.17) ...........190

Appendix E9: Viable Cell Count in Induced (42 °C) and Uninduced Cultures Supplemented with 5.0 g/l Glucose and 5.0 g/l Glycerol (Figure 4.18, 4.19) ..............................................................191

Appendix E10: pH Profile in Cultures Supplemented with 5.0 g/l Glucose and 5.0 g/l Glycerol (Figure 4.20) ......................................................192

Appendix E11: Glucose Consumption Profile in Induced (42 °C) and Uninduced Cultures (Figure 4.21) ......................................................193

Appendix E12: Effect of Acetic Acid and pH on Growth and Protein Expression (Figure 4.22, 4.23, and 4.24) .................................194

Appendix E13: Dry Biomass Concentration and β-galactosidase Activity of Cultures Induced at 38 °C, 40 °C and 42 °C (Figure 4.30, 4.31, 4.33, 4.34, 4.35, 4.36, and 4.37) ......................................................195

Appendix E14: Viable Cell Counts of Cultures Instantaneously Induced at 38 °C, 40 °C and 42 °C (Figure 4.32) .................................196

Appendix E15: Dry Biomass Concentration, Volumetric and Specific β-galactosidase Activities from Ramped-temperature Upshift Experiment (Figure 4.38) ......................................................197

APPENDIX F: Simulation Results

Appendix F1: Simulation Data of Initial Number of Active Plasmid per Cell (n) at Different Temperatures (Figure 5.2) ......................198

Appendix F2: Simulation Data of Number of Active Plasmid per Cell
Appendix F3: Simulation Data of Percentage of Kanamycin-resistant Cells at 38 °C, 40 °C, and 42 °C (Figure 5.4) .........................200

Appendix F4: Simulation Data of Volumetric β-galactosidase activity (U/ml) at 38 °C, 40 °C and 42 °C (Figure 5.5) ...............201

Appendix F5: Simulation Data of Dry Cell Biomass Concentration (g/l) at 38 °C, 40 °C, and 42 °C (Figure 5.6) .........................202

Appendix F6: Simulation Data of Specific β-galactosidase Activity (U/mg) at 38 °C, 40 °C, and 42 °C (Figure 5.7) .........................203

Appendix F7: Test of Model for Goodness of Fit on Observed and Predicted Data of Specific Expression (U/mg) at 38 °C, 40 °C, and 42 °C ..................................................204

Appendix F8: Simulation Data of Dry Cell Biomass Concentration, Volumetric and Specific β-galactosidase Activities from Ramped-temperature Upshift Experiment (Figure 5.8, 5.9 and 5.10) ................................................................205

Appendix F9: Test of Model for Goodness of Fit on Observed and Predicted Data of Specific Expression (U/mg) of Ramped Upshift Experiment .........................................................206
NOMENCLATURE

c Rate constant relating protein expression to plasmid induction (Conversion factor)

E_a Specific Activation Energy (kJ kmol⁻¹)

F_{cell} fraction of kanamycin-resistant cell (dimensionless)

k temperature ramp constant (°C per hour)

k parameter or reaction rate constant

k_d specific death rate of viable cell (hour⁻¹)

k_{lys} specific rate of cell lysis (hour⁻¹)

k_{prot} denaturation rate of protein by protease (hour⁻¹)

n number of active plasmid per cell (dimensionless)

P Volumetric β-galactosidase activity (U ml⁻¹)

p probability of a given plasmid (out of ‘n’ plasmid) is still active (dimensionless)

q probability of a given plasmid (out of ‘n’ plasmid) is inactive (dimensionless)

R Gas-law Constant (8.314 kJ kmol⁻¹K⁻¹)

t induction time

T temperature (°C)

T_K Absolute Temperature (Kelvin)

T_{max} maximum temperature (°C)

X dry biomass concentration (g l⁻¹)

X_{max} maximum dry biomass concentration (g l⁻¹)

X_v viable cell concentration (g l⁻¹)
Greek Letters

δ unobserved variable

λ plasmid decay constant (\text{hour}^{-1})

μ specific growth rate (\text{hour}^{-1})

μ_{\text{max}} maximum specific growth rate (\text{hour}^{-1})

θ parameter estimates (value)
CHAPTER ONE

INTRODUCTION

The advancement of post-genomic biotechnology in the past two decades has increased heterologous protein titer in fermentations many folds. A great deal of effort has been spent on developing better expression hosts, vectors, as well as culture environments. As a relatively new subject, modern biotechnology is always being regarded as a ‘marriage of convenience’ between two fields, bioprocess engineering and genetic engineering. As far as bioprocess engineering is concerned, the progress in this discipline has not been as dynamic as in its counterpart, genetic engineering. The evidence on this matter is clearly seen if the number of research publications are used as a yardstick. Although the traditional focus of bioprocess engineering was solving classical engineering problems such as mass and energy transfer issues in the scaling up of bioreactors, recently applications have been extended into more specialized areas in fermentation such as regulating heterologous protein expression.

The interdisciplinary nature of the tasks in managing recombinant protein expression from recombinant organisms needs wider and deeper understanding at various aspects including the cell’s physiological demands, genetic stability, as well as fermentation strategies. In most cases, the interactions between these aspects are unavoidably simplified by several empirical relationships. Unfortunately, due to a fairly large variation and complexity of the molecular biology among the recombinant organisms, the generality of the optimization and characterization approach for one
system may not necessarily be extendable to another system. Thus, a robust and flexible bioprocess approach may be the best tool.

1.1 *Escherichia coli* as an Expression Host

The choice of expression hosts for heterologous protein production in the fermentation industry is determined by several factors such as processing strategies, flexibility in gene manipulation, genetic stability, promoter and vector effectiveness, regulatory approval, and patent considerations. Among the popular expression hosts for the heterologous proteins are *Escherichia coli*, *Bacillus subtilis*, Yeasts (*Saccharomyces cerevisiae, Kluyveromyces lactis, Pichia pastoris, Yarrowia lipolytica*), Streptomyces, filamentous fungi, insect cells, and mammalian cells (Shuler and Kargi, 1992).

In comparison to other expression hosts, *Escherichia coli* remains the most frequently employed workhorse in the production of heterologous proteins. *Escherichia coli* enjoys GRAS (generally recognized as safe) status, which greatly simplifies the experimentation required for regulatory approval and clinical trials. The popularity of *E.coli* as the expression host is due to the extensive knowledge of its genetics and physiology. This extensive knowledge in *E.coli*’s sub-cellular biology largely facilitates the manipulation of genetics including the development of well-characterized vectors and promoters (Yarranton, 1992). A depth of understanding of its physiology such as the Crabtree-like effect promotes the development of feeding strategies and allows minimizing the fermentative condition in *Escherichia coli* cultures. As a result, high-cell density cultures of up to 80 g/l dry weight (Lee and Chang, 1993; Mori *et al.*, 1979) is
possible to achieve. Other attractive features of *E. coli* including high growth rate, and growth in inexpensive and easily available medium, also significantly contributed to its popularity as the expression host.

However, traditionally *E. coli* does not excrete the expressed protein into the culture broth. Hence, the disruption of cells to release the intracellular products is inevitable. The disruption of *E. coli* cells may release pyrogens (or endotoxins) that burden the purification steps. Nevertheless, endeavors to improve the secretion and excretion level of heterologous protein to the periplasmic and culture broth respectively have become an attractive research subject (e.g. in Blight *et al.*, 1994; Weikert and Bailey, 1998). As a prokaryote, *E. coli* has some disadvantages to express the cloned eucaryotic genes correctly due to lack of some biosynthetic mechanisms such as post-translational modifications, and mRNA splicing.

1.2 Studies of Heterologous Protein Expression Using Regulatable Strong Promoters in *Escherichia coli* Cultivations

As opposed to the constitutive (continuous or uncontrolled) systems, induction allows the separation of a production phase from the growth phase. Hence, high heterologous protein titer following the induction of high-density culture can be realized. The application of induction is especially critical if the expressed protein is toxic to the host. One of the basic features of an inducible system for heterologous protein expression is the use of a strong regulatable promoter. Strong promoters have a high affinity for RNA polymerase. Hence, once the inducible system is exposed to the
inducing agents, the transcription of gene or operon adjacent to the promoter is enhanced. The disengagement of RNA polymerase from the plasmid after reading through the operon is also crucial, or otherwise the transcription of the undesired genes may happen. Poor disengagement of RNA polymerase from the plasmid may also potentially interrupt the plasmid-copy-number control mechanism, which may lead to the runaway replication of the plasmid and the death of the host.

Some of the most widely used regulatable strong promoters in *Escherichia coli* are *lac*, *tac* (a variation of *lac*, induced by a gratuitous inducer, isopropyl-β-D-thiogalactopyranoside (IPTG)), *trp*, *pL*, and T7. Among these promoters, *pL* and *tac* (or *lac*) promoter are most commonly employed. This is probably due to the extensive understanding of the promoter characteristics on the molecular level. In the lambda *pL/cI857* promoter system, the protein expression from the cloned gene(s) on the plasmid DNA is suppressed by lowering the culture temperature, for example between 28 °C and 32 °C. At these temperatures, the binding of the repressor protein cI857 blocks the transcription of gene(s) downstream of the *pL* promoter. The inactivation, and hence the detachment of repressor protein cI857 from the plasmid DNA occurs when the temperature is at 38 °C or above. Under some circumstances, the *pL* promoter is favored over the *tac* promoter due to cost and safety considerations. The use of a large amount of IPTG to activate the *tac* promoter system is not only a problem in downstream processing, but it is also costly and toxic to humans. However, it may cause induction more rapidly and accurately than the *pL* promoter system in large fermentation vessels. This is due to the rapid delivery of IPTG in comparison to heat in the fermentation broth. The thermal lag
in a larger scale and the need of suboptimal pre-induction temperatures in the \( p^L \) promoter system may also significantly prolong the culture time.

Since the development of \( p^L/\text{cl}_{857} \) expression system on the cloning vector in the late seventies (Bernard and Helsinki, 1979; Bernard et al., 1979), many improvements on the heterologous protein expression using bioprocess engineering knowledge on this system has been reported. Table 1.1 provides some overview of the bioprocess studies of \( p^L \) promoter system performed in *Escherichia coli* cultures. It is evident from the table that the focus of the studies was on three major areas, namely the importance of specific growth rate (through fed-batch or continuous culture) on productivity, the intracellular stability and solubility of expressed proteins, and optimization of induction temperature. Although some temperature-dependent modeling on the \( p^L \) promoter system has been reported (e.g. in Hørtacsu and Ryu, 1990; Cockshott and Bogle, 1992, 1999), the direct effect of temperature on the productivity has not been described.

1.3 Problems Associated with the \( p^L \) Promoter System and Protein Overexpression in *Escherichia coli*

Although the \( p^L \) promoter system allows high level expression of heterologous proteins in *Escherichia coli* cultures, the deleterious side effects of protein overproduction on the host and the expressed proteins have always been a major concern. Among the obvious problems are the formation of inclusion bodies of the expressed proteins, the degradation of protein by the stress-induced proteolysis, the increased secretion of acetic acid following the induction, and plasmid instability.
Table 1.1: Studies employing λp1/cI857 system for protein expression in *Escherichia coli*

<table>
<thead>
<tr>
<th>Study</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect of temperature on plasmid instability</td>
<td>Siegel &amp; Ryu (1985)</td>
</tr>
<tr>
<td>Heat shock effect on product degradation</td>
<td>Goff <em>et al.</em> (1984)</td>
</tr>
<tr>
<td>Pre-induction μ in fed-batch study</td>
<td>Zabriskie <em>et al.</em> (1987)</td>
</tr>
<tr>
<td>Timing and duration of induction</td>
<td>diPasquantonio <em>et al.</em> (1987)</td>
</tr>
<tr>
<td>Effect of substrate feeding on productivity</td>
<td>Whitney <em>et al.</em> (1988)</td>
</tr>
<tr>
<td>Optimization of temperature and duration</td>
<td>Okita <em>et al.</em> (1989)</td>
</tr>
<tr>
<td>Pre-induction μ in chemostat</td>
<td>Curless <em>et al.</em> (1990)</td>
</tr>
<tr>
<td>Production modelling in fed-batch culture</td>
<td>Cockshott &amp; Bogle (1992)</td>
</tr>
<tr>
<td>Genetically stuctured modelling in chemostat study</td>
<td>Hortacsu &amp; Ryu (1990)</td>
</tr>
<tr>
<td>Comparison between batch and fed-batch cultures</td>
<td>Strandberg &amp; Enfors (1991)</td>
</tr>
<tr>
<td>1000-L fermentation and downstream study</td>
<td>Strandberg <em>et al.</em> (1991)</td>
</tr>
<tr>
<td>Fed-Batch with DO-Stat and induction timing</td>
<td>Seo <em>et al.</em> (1992)</td>
</tr>
<tr>
<td>Effect of pH and timing of induction</td>
<td>Kamasawa <em>et al.</em> (1994)</td>
</tr>
<tr>
<td>Effect of pH and temperature</td>
<td>Horiuchi <em>et al.</em> (1994)</td>
</tr>
<tr>
<td>Pre-induction μ in chemostat study</td>
<td>Curless <em>et al.</em> (1994)</td>
</tr>
<tr>
<td>Comparison between Temp. and IPTG induction</td>
<td>Seeger <em>et al.</em> (1995)</td>
</tr>
<tr>
<td>Effect of energy source on protein degradation</td>
<td>Yang &amp; Enfors (1995)</td>
</tr>
<tr>
<td>Study of various temperatures on expression</td>
<td>Yabuta <em>et al.</em> (1995)</td>
</tr>
<tr>
<td>Pre-induction μ</td>
<td>Ryan <em>et al.</em> (1996)</td>
</tr>
<tr>
<td>Profiled temperature upshift</td>
<td>Lamotte <em>et al.</em> (1996)</td>
</tr>
<tr>
<td>Production modelling in fed-batch culture</td>
<td>Cockshott &amp; Bogle (1999)</td>
</tr>
</tbody>
</table>

The overexpression of a foreign protein in *Escherichia coli* regulated by a strong promoter, particularly those obtained from eucaryotes, has often been followed by the formation of inclusion bodies. These bodies are insoluble and biologically inactive resulting from the misfolding of the protein. However, under some circumstances, the formation of these bodies is favorable and intentionally designed since they are less prone to proteolytic attack. Nevertheless, time-consuming steps, such as solubilization and renaturation to recover the bioactive proteins are necessary.
The upshift of culture temperature and the following overexpression of heterologous proteins force the host to make a physiological adjustment. Amongst the important adjustments is the heat-shock response. This event is a universal response of organisms to the environmental stress such as temperature, nutrient starvation and osmotic pressure. The term ‘heat-shock’ was coined since it was first observed in the super-optimal temperature shifts (Enfors, 1992). This response is essential to the survival of the cells in harsh environments. Unfortunately, one of the main attributes of the heat shock response is the induction of the stress-related proteolysis that seriously degrades the heterologous proteins (Goff et al., 1988).

The accumulation of acetic acid from the Crabtree-like effect in *Escherichia coli* cultures is a well-characterized phenomenon. This effect is usually overcome by maintaining the growth-limited substrate (e.g. glucose or glycerol) concentration below the threshold levels. However, in the event of energy-intensive activities such as protein overexpression, the release of acetic acid to the culture medium is evidently increased. This is due to the limitation of oxidative metabolism resulting from insufficient turnover of ATP and reducing agents such as NADH and NADPH$_2$ (Han et al., 1992). As the alternative, the metabolism of glucose switches to fermentative pathway in order to generate more instant energy. This would implicate a feedback repression on the tricarboxylic acid (TCA) cycle, and hence protein synthesis. The acetic acid interrupts protein expression and growth mechanisms.
Although the problem with plasmid instability in *E. coli* is not as serious as in *B. subtilis* (Shuler and Kargi, 1992) or yeast (Zhang 1997), this may be a consequence of protein overexpression. A study by Stueber and Bujard (1982) revealed that the transcription of protein from efficient promoters could interfere with plasmid replication and diminish protein expression. As far as the $p^\perp$ promoter system is concerned, there has been no particular bioprocess techniques designed for the purpose of minimizing plasmid instability upon a temperature upshift. Also, the effect of induction temperature on plasmid stability has not been described previously. Among the reported bioprocess strategies to circumvent the plasmid instability problem are listed in Table 1.2. Studies indicated that the propagation of plasmid-free cells might be overcome by limiting their growth rate by two-stage culture or recycling, separation of plasmid-free segregants, and immobilization. Regulation of specific growth rate through two-stage fermentation and recycling either at different dilution rate or different substrate successfully minimized the destabilizing effect on plasmid. The unproductive sub-population may be separated by using selective flocculation or aqueous two-phase system. Although this may eventually increase the final specific productivity, these approaches may be problematic and costly at large scale. The employment of cell immobilization is mainly due to the fact that the restriction of growth may minimize the growth-related segregational loss. Immobilization of cells was also applied in order to 1) maintain high cell density, so that continuous operation without relying on cell growth can be applied, 2) facilitate cell separation, 3) obtain higher yield by enabling the use of mixed and spatially localized microbial cultures (Dulieu et al., 1999)
From the discussion and the review of earlier studies, it could be summarized that the protein expression in the $p^L$ promoter system is profoundly determined by the culture environment, the cell physiology, and the interaction between them. In comparison to other inducible methods, the $p^L$ promoter system is unique due to the twofold role of temperature; it is both an inducing agent and a key factor in all biological reactions. In view of this, it is indeed a great advantage to develop an efficient integrated approach for describing the direct influence of temperature on protein expression system in the $p^L$ promoter system.

### Table 1.2: Bioprocess strategies in minimizing plasmid segregational instability

<table>
<thead>
<tr>
<th>Inhibition of growth Advantage of plasmid-free Cell</th>
<th>Two-stage cultivation</th>
<th>Siegel and Ryu (1985)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recycling between:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) different dilution rates</td>
<td>Impoolsup et al. (1989)</td>
<td></td>
</tr>
<tr>
<td>ii) different substrate concentration</td>
<td>Stephen &amp; Lyberatos (1985)</td>
<td></td>
</tr>
<tr>
<td>Separation of plasmid-free cells from plasmid-bearing cells</td>
<td>Selective flocculation</td>
<td>Henry et al. (1990)</td>
</tr>
<tr>
<td>Aqueous two-phase system</td>
<td>LaMarca et al. (1990)</td>
<td></td>
</tr>
<tr>
<td>Whole-cell immobilization</td>
<td>on <em>E. coli</em></td>
<td>Inloes et al. (1983)</td>
</tr>
<tr>
<td>on <em>B. subtilis</em></td>
<td>Mosbach et al. (1983)</td>
<td></td>
</tr>
</tbody>
</table>

Adaptation from Kumar et al. (1991)

### 1.4 Objectives

In the present study, a recombinant *Escherichia coli* that bears a multiple copy of plasmid pBRG 401 is used as a model system. The overexpression of β-galactosidase from the *lac Z* gene, which was isolated from the chromosomal genome, and encoded on the plasmid was achieved by upshifting the culture temperature between 38 °C and 42 °C. The plasmid also carried the *kan* gene to give the host a resistance towards kanamycin.
The selective pressure was not included in the production phase, so that the effect of temperature shift on plasmid instability could be investigated. This was appropriate for the study of the \( p^L \) promoter system for the following reasons. An unstable plasmid would give an opportunity to investigate factors that contribute to its instability. The availability of an inexpensive, fast and reliable assay of \( \beta \)-galactosidase would facilitate product detection. The stability of \( E.coli \) \( \beta \)-galactosidase has also been well characterized (e.g. in Banzet et al., 1992) and this would allow the bioactivity to be retained in a reasonable range of operating conditions. The universality of \( \beta \)-galactosidase in kinetic studies in other hosts such as \( Saccharomyces \) \( cerevisiae \) (Hadjito \textit{et al.}, 1992) and \( Bacillus \) \( subtilis \) (Park \textit{et al.}, 1992) may be also important for comparison purposes. Further details of the characteristics of \( E.coli \)'s \( \beta \)-galactosidase are summarized in Appendix A.

The hypotheses underlying this thesis are:

1) Standard conditions employed for thermal induction are not optimal

2) The loss of plasmid expression in thermal induction system is not adequately described

To test these hypotheses, the objectives of this study were as follow:

1) to investigate and identify the influential factors that determine the protein expression using the \( p^L/cI_{857} \) system

2) to introduce an integrated approach in describing the effect of the inducing agent (temperature) on plasmid stability, cell growth, and product expression