FERMENTATION AND RECOVERY OF *Eco*RI ENDONUCLEASE USING TWO DIFFERENT GENETICALLY ENGINEERED *E.coli* STRAINS

by

Candan (Tamerler) Yıldır

B.S. in Ch.E., Boğaziçi University, 1989

M.S. in Ch.E., Boğaziçi University, 1991

Submitted to the Institute for Graduate Studies in Science and Engineering in partial fulfillment of the requirements for the degree of

Doctor

of

Philosophy

Boğaziçi University

1997

ACKNOWLEDGMENTS

The research presented in this thesis was conducted at the Chemical Engineering Department of Boğaziçi University and was supported by the Turkish State Planning Organization (DPT) Technological Research Projects No. 91K120460 and 95K120320.

I would like to express my sincere gratitude to my thesis supervisors Prof. Dr. Z. İlsen Önsan and Prof. Dr. Betül Kırdar for their continuous guidance, encouragement, motivation and support. It was a unique opportunity for me to work with them during my M.S. and Ph.D studies. They have created such a wonderful laboratory in which it was a pleasure to work.

I would like to thank Prof. Dr. Amable Hortaçsu, Prof. Dr. Nejat Dalay for the time they have devoted to reading and commenting on my thesis.

I am indebted to Gülnur Birol, Nilgün Akın and A. Erhan Aksoylu for sharing all the good and bad moments and for their everlasting support whenever I needed it.

I wish to express my special thanks to Mustafa Altınkaya of the EE for his guidance and support during the final stages of my thesis work.

Special thanks are due to Bilgi Dedeoğlu and Mine Yediç of the ChE Department for their kind help and assistance during the entire period of my studies.

Heartfelt thanks are due to all my friends at ChE Department and especially to all the members of the Biotechnology group (past ! and present) who made the hardworking days more than enjoyable.

I am especially grateful to my brother, Kudret Tamerler who has always encouraged and supported me.

To my Mom and Dad: I thank you from the bottom of my heart for your endless love and support.

Surely, I am deeply indebted to my husband, Ümit Yıldır. This work would not have been possible without his love, unlimited patience and understanding.

A special dedication to a true miracle: Emre Tamerler , my nephew who has come into the world and enriched our lives.

ABSTRACT

A laboratory scale procedure developed for the purification of *Eco*RI restriction endonuclease has been applied to two different overproducing *Escherichia coli* strains. The yields obtained at the final stage of the purification are $1.3x10^5$ U/g cells for *E.coli* M5248 and $3.3x10^6$ U/g cells for *E.coli* 294. Induction parameters were optimized in order to increase the yield of the *Eco*RI endonuclease isolated from both recombinant *E.coli* strains. For the *E.coli* 294 (pPG430) strain, 0.1 mM IPTG concentration at an optical density of 1.2 at 595 nm over an induction period of 6 hours were determined to be the optimum conditions for induction. For the *E.coli* M5248 (pSCC2) strain, the induction of *Eco*RI endonuclease was achieved by a temperature-shift from 32°C to 42 °C at an optical density of 1.0 at 590 nm over an induction period of 5 hours.

Investigation of the stability of plasmids pPG430 and pSCC2 showed that in the case of *E.coli* 294 (pPG430), minimal medium results in a decrease in enzyme productivity and plasmid stability, and in the case of *E.coli* M5248 (pSCC2) cells, temperature-upshift leads to a decrease in the percentage of plasmid-containing cells and in *Eco*RI productivity.

The unstructured modeling of cell growth has shown that substrate-inhibited kinetics is successful in predicting the growth behavior of both *E.coli* 294 and *E.coli* M5248. A three-compartment model was developed to describe the dynamic changes in the intracellular components of *E.coli* 294 cells. The model gives a good description of the growth kinetics with an initial glucose concentration of 10 g/L.

ÖZET

*Eco*RI restriksiyon endonükleazını saflaştırmak için geliştirilen ufak ölçekli saflaştırma yöntemi iki farklı rekombinant *Esherichia coli* suşuna uygulanmıştır. Saflaştırma sonucunda *E.coli* M5248 için $1.3x10^5$ U/g-hücre ve *E.coli* 294 için $3.3x10^6$ U/g-hücre verim elde edilmiştir. *Eco*RI endonükleazının her iki rekombinant *E.coli* suşundan üretim verimini arttırmak amacıyla indüklenme parametreleri incelenmiştir. *E.coli* 294 (pPG430) suşundan 595 nm dalga boyunda 1.2 optik yoğunluğa erişildiğinde eklenen 0.1 mM IPTG ile 6 saatlik indüklenmenin enzim üretimi için optimum koşulları sağladığı saptanmıştır. *E.coli* M5248 (pSCC2) suşundan *Eco*RI eldesi için optimum koşullar 590 nm dalga boyunda 1.0 optik yoğunluğa erişildiğinde fermentasyon sıcaklığını 32°C den 42°C ye yükselterek 5 saat süre ile indüklenmesi olarak bulunmuştur.

Plasmid pPG430 ve pSCC2'nin kararlığına ilişkin çalışmalar, *E.coli* 294 (pPG430) suşu için, minimal besi ortamında büyütüldüğünde enzim üretkenliğinin ve plasmid kararlılığının azaldığını, *E.coli* M5248 (pSCC2) suşu için ise, sıcaklık artışının plasmid kararlılığının ve *Eco*RI üretkenliğinin azalmasına neden olduğunu göstermiştir.

E.coli 294 ve *E.coli* M5248 hücrelerinin büyüme kinetiğinin yapısal olmayan sübstrat-kısıtlaması içeren modeller ile tanımlanabildiği belirlenmiştir. 3-kademeli bir yapısal model kullanılarak *E.coli* 294 hücrelerinin hücre içi bileşenlerinin değişimini tanımlayan bir kinetik model geliştirilmiş ve 10g/L başlangıç glikoz derişiminde elde edilen deneysel verilere uygunluğu saptanmıştır.

TABLE OF CONTENTS

page

ACKNOWLEDGMENT iii
ABSTRACT v
ÖZET vi
LIST OF FIGURES xii
LIST OF TABLES xxi
LIST OF SYMBOLS xxvi
1. INTRODUCTION 1
2. THEORETICAL BACKGROUND 3
2.1 Restriction Endonucleases 3
2.1.1 General Information on Restriction Endonuclease 3
2.1.2 Structure and Function of Type II Restriction Enzymes 12
2.1.3 Purification of Restriction Enzymes 16
2.1.1.1 Sources and Growth Conditions 17
2.1.1.2 Preparation of the Cell Extract 19
2.1.1.3 Chromatographic Techniques 22
2.1.1.4 Enzyme Assays 25
2.1.1.5 Criteria of Purity 28
2.1.1.6 Stability and Storage of Restriction Endonucleases 28
2.2 Purification of *Eco*RI Endonuclease 29
2.3 Recombinant Fermentation Processes 33
2.3.1 Bioprocess Design Variables 33
2.3.2 Optimization of Recombinant Fermentation Processes 37
2.3.2.1 Temperature 38
2.3.2.2Time of Induction 38
2.3.2.3 Oxygen Demand and Dilution Rate 39

2.3.3.4 Stability of Recombinant Plasmids 39
2.4 Modeling of Bacterial Cell Growth and Product Foormation 42
2.4.1 Unstructured Models 43
2.4.1.1 Inhibition-Free Single-Substrate Limiting Kinetics 44
2.4.1.2 Substrate-Independent Growth Kinetics 45
2.4.1.3 Substrate and Product Inhibition Kinetics 46
2.4.2 Structured Models 48
2.4.3 Segregated Models 52
3. MATERIALS AND METHODS 55
3.1 Materials 55
3.1.1 Bacterial Strains and Plasmids 55
3.1.2 Bacterial Culture Media 55
3.1.3 Chemicals
3.1.3.1 DNA Size Marker Used in the Electrophoresis 57
3.1.3.2 SDS Molecular Weight Markers 57
3.1.4 Buffers and Solutions 58
3.1.4.1 Assay Buffer 58
3.1.4.2 Lysis Buffer 58
3.1.4.3 Dialysis Buffer 59
3.1.4.4 Storage Buffer 59
3.1.4.5 Agarose Gel Electrophoresis Buffers 59
3.1.4.6 Polyacrylamide Gel Electrophoresis Buffers 59
3.1.4.7 Buffers and Solutions used in Plasmid DNA Isolation 61
3.1.4.8 Solutions used in RNA Estimation 63
3.1.5 Laboratory Equipments 63
3.2 Experimental Methods 66
3.2.1 Preparation of Pre-culture 66
3.2.2 Large Scale Plasmid Isolation by Alkaline-Lysis Method 66

3.2.3 Transformation of *E.coli M5248* cells by Plasmid pSCC2 by $CaCl_2$ Method 68

3.2.4 Cracking Procedure for Rapid Plasmid Detection in *E.coli M5248* 69

3.2.5 Determination of the Population Growth Curve 69

3.2.6 Harvesting Cells 70

3.2.7 Purification of the Enzyme 70

3.2.7.1Preparation of the Cell Lysate 70

3.2.7.2 Application of Phosphocellulose Chromatography 71

3.2.7.3 Application of Hydroxyapatite Chromatography 72

3.2.7.4 Application of Second Phosphocellulose Chromatography 73

3.2.7.5 Storage 73

3.2.8 Application of Quality and Control Tests 73

3.2.8.1 Cut-Ligate-Recut Test 73

3.2.8.2 Overdigestion Test 74

3.2.8.3 Determination of the purity of the enzyme 74

3.3 Analysis of Fermentation 74

3.3.1 Determination of Dry Cell Weight 74

3.3.2 Determination of Substrate Concentration 75

3.3.3 Determination of Protein Content 76

3.3.4 Determination of *Eco*RI Endonuclease Activity 77

3.3.5 Determination of Intracellular RNA Content by Orcinol Method 78

3.3.6 Determination of Plasmid Content 79

3.3.7 Determination of Total Viable Cells 80

3.3.8 Determination of the Percentage of the Plasmid Containing Cells 81

4. RESULTS AND DISCUSSIONS 82

4.1 Optimization of induction parameters for the production of *Eco*RI endonuclease by *E.coli* 294(pPG430) 84

4.2 Optimization of the temperature-shift period for the production of *Eco*RI endonuclease by *E.coli* M5248(pSCC2) 89

4.3 Purification of *Eco*RI restriction endonuclease from two different recombinant *E.coli* strains 93

4.4 Investigation of the plasmid stabilities of *E.coli* 294(pPG430) and *E.coli* M5248(pSCC2) 106

4.4.1Stability of plasmid pSCC2 in *E.coli* M5248 cells 107

4.4.1.1 The effect of the temperature shift on plasmid stability 107

4.4.1.2 The effect of cycling time on product yield 115

4.4.2 Stability of plasmid pPG430 in *E.coli* 294 cells 118

4.4.3 Investigation of the stabilities of plasmids pPG430 and pSCC2 after successive generations 124

4.4.4 Mathematical methods in the analysis of plasmid stability 126

4.5 Effect of Substrate Concentration on the Growth and *Eco*RI Production of Recombinant *E.coli* 294 and *E.coli* M5248 strains 134

4.6 Efffect of Glucose Concentration on the Intracellular Components of Recombinant *E.coli* 294 Strain 146

4.7 Investigation of copy numbers of plasmids pPG430 and pSCC2 152

4.8 Growth Kinetics by Using Unstructured Models 156

4.8.1 Overall Yield Factors 158

4.8.2 Maximum specific growth rates 161

4.9 Determination of Kinetic Parameters by Using Structured Models 164

5. CONCLUSIONS AND RECOMMENDATIONS 171

5.1 Conclusions 171

5.2 Recommendations 174

APPENDIX A. RELATION BETWEEN OPTICAL DENSITY AND DRY CELL WEIGHT 175

APPENDIX B. CALIBRATION CHART FOR BRADFORD DYE BINDING METHOD 177

APPENDIX C. EXPERIMENTAL DATA FOR BIOMASS AND SUBSTRATE CONCENTRATION 178

APPENDIX D. EXPERIMENTAL DATA FOR CONCENTRATION OF INTRACELLULAR COMPONENTS 182

APPENDIX E. CURVE FITTING TECHNIQUE AND COMPUTER PROGRAM 190

APPENDIX F. COMPUTER PROGRAM USED IN UNSTRUCTURED MODELING 200

APPENDIX G. COMPUTER PROGRAM USED IN 3-COMPARTMENT STRUCTURED MODELING 208

REFERENCES 219

LIST OF FIGURES

		Page
FIGURE 2.1.	Restriction-modification.	4
FIGURE 2.2.	*Eco*RI restriction enzyme's sticky ended cut	8
FIGURE 2.3.	Stereo drawings of Cα atoms of *Eco*RI endonuclease and the substrate DNA	14
FIGURE 2.4.	The steps involved in the crude extract preparation for purification of intracellular restriction endonuclease from microorganisms	21
FIGURE 2.5.	Schematic diagram of the in fluence of genetic and environmental factors on overall product synthesis	34
FIGURE 2.6.	Different perspective for cell population kinetic representations	43
FIGURE 4.1.	Structure of plasmid pPG430	85
FIGURE 4.2.	Growth curve of *E.coli* 294 (pPG430)	86
FIGURE 4.3.	The effect of the period of induction on the *Eco*RI specific activity	87
FIGURE 4.4.	Structure of plasmid pSCC2	90
FIGURE 4.5.	Growth curve for *E.coli M5248* cells (pSCC2) at 32 °C	90

FIGURE 4.6. The effects of temperature-shift on the yield and the specific activity of *Eco*RI endonuclease (*E.coli* M5248) 92

FIGURE 4.7. The effect of temperature shift on the recombinant plasmid stability of *E.coli* M5248. 93

FIGURE 4.8. Determination of the sonication period for *E.coli* 294 (pPG430) cells 95

FIGURE 4.9. Determination of the sonication period for *E.coli* M5248 (pSCC2) cells 95

FIGURE 4.10. Elution profile of the first phosphocellulose column chromatography for *E.coli* 294 strain 98

FIGURE 4.11. Elution profile of the first phosphocellulose column chromatography for *E.coli* M5248 strain 99

FIGURE 4.12. Elution profile of the second phosphocellulose column chromatography for *E.coli* M5248 strain 100

FIGURE 4.13. SDS-PAGE analysis of the protein content 102

FIGURE 4.14. Overdigestion assay of purified *Eco*RI endonuclease 103

FIGURE 4.15. Cut-Ligate-Recut test of purified *Eco*RI endonuclease 105

FIGURE 4.16. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 1 hour temperature-shift 108

FIGURE 4.17. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 2 hour temperature-shift 108

FIGURE 4.18. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 3 hour temperature-shift 109

FIGURE 4.19. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 4 hour temperature-shift 109

FIGURE 4.20. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 5 hour temperature-shift 110

FIGURE 4.21. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 6 hour temperature-shift 110

FIGURE 4.22. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 7 hour temperature-shift 111

FIGURE 4.23. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 8 hour temperature-shift 111

FIGURE 4.24. The change in recombinant viable cell counts with respect to time in batch cultures exposed to different temperature-shift periods between 1-8 hours 113

FIGURE 4.25. The change in the percentages of plasmid containing cells with respect to time in batch cultures exposed to different temperature-shift periods between 1-8 hours 114

FIGURE 4.26. The change in the specific growth rates (μ) and *Eco*RI endonuclease production during T-shift period of 8 hours 115

FIGURE 4.27. Growth behaviour of the recombinant cultures which were shifted from 32°C to 42°C between 1-5 hours, then shifted back to 32°C 117

FIGURE 4.28. The growth curves of host and recombinant *E.coli* 294 cells in selective and non-selective complex media 119

FIGURE 4.29. The growth curves of host and recombinant *E.coli* 294 cultures grown in M9 medium 121

FIGURE 4.30. The effect of medium composition on the percentage of plasmid containing cells 123

FIGURE 4.31. The variation in the stabilities of plasmid pSCC2 in *E.coli* M5248 and pPG430 in *E.coli* 294 during 20 successive batches 125

FIGURE 4.32. CaseI: lnp- versus time plot for *E.coli* M5248 exposed to 3 hours T-shift 130

FIGURE 4.33. CaseI: lnp- versus time plot for *E.coli* M5248 exposed to 5 hours T-shift 131

FIGURE 4.34. CaseI: lnp- versus time plot for *E.coli* 294 cultured in M9 LB media 131

FIGURE 4.35. CaseII: lnp_+ versus time plot for *E.coli* M5248 exposed to 3 hours T-shift 132

FIGURE 4.36. CaseII: lnp_+ versus time plot for *E.coli* M5248 exposed to 5 hours T-shift 132

FIGURE 4.37. CaseII: lnp- versus time plot for *E.coli* 294 cultured in M9 and LB media 133

FIGURE 4.38. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 1 g/L) 136

FIGURE 4.39. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 5 g/L) 137

FIGURE 4.40. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 10 g/L) 137

FIGURE 4.41. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 15 g/L) 138

FIGURE 4.42. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 30 g/L) 138

FIGURE 4.43. Time profiles of cell mass and residual glucose concentration for *E.coli* M5248 cells in minimal medium (S_0 = 1 g/L) 139

FIGURE 4.44. Time profiles of cell mass and residual glucose concentration for *E.coli* M5248 cells in minimal medium (S_0 = 5 g/L) 139

FIGURE 4.45. Time profiles of cell mass and residual glucose concentration for *E.coli* M5248 cells in minimal medium (S_0 = 10 g/L) 140

FIGURE 4.46. Time profiles of cell mass and residual glucose concentration for *E.coli* M5248 cells in minimal medium (S_0 = 15 g/L) 140

FIGURE 4.47. Time profiles of cell mass and residual glucose concentration for *E.coli* M5248 cells in minimal medium (S_0 = 30 g/L) 141

FIGURE 4.48. Maximum specific growth rates at different glucose concentrations for the recombinant strains *E.coli* M5248 and *E.coli* 294 142

FIGURE 4.49. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 1 g/L) 143

FIGURE 4.50. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 5 g/L) 144

FIGURE 4.51. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 10 g/L) 144

FIGURE 4.52. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 15 g/L) 145

FIGURE 4.53. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 30 g/L) 145

FIGURE 4.54. The effect of initial glucose concentration on the total activity of the *Eco*RI endonuclease for both of the recombinant *E.coli* strains 146

FIGURE 4.55. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium with initial glucose concentration of 20 g/L 147

FIGURE 4.56. Time profiles of RNA, plasmid content and total protein content for 30 g/L initial glucose concentration (*E.coli* 294) 149

FIGURE 4.57. Time profiles of RNA, plasmid content and total protein 149

content for 20 g/L initial glucose concentration (*E.coli* 294)

FIGURE 4.58. Time profiles of RNA, plasmid content and total protein content for 15 g/L initial glucose concentration (*E.coli* 294) 150

FIGURE 4.59. Time profiles of RNA, plasmid content and total protein content for 10 g/L initial glucose concentration (*E.coli* 294) 150

FIGURE 4.60. Time profiles of RNA, plasmid content and total protein content for 5 g/L initial glucose concentration (*E.coli* 294) 151

FIGURE 4.61. Time profiles of RNA, plasmid content and total protein content for 1 g/L initial glucose concentration (*E.coli* 294) 151

FIGURE 4.62. Time profile of the total *Eco*RI activity in *E.coli* 294 cell cultures grown in minimal medium (S_0 = 20 g/L) 152

FIGURE 4.63. The effect of the number of generations on the plasmid copy number (*E.coli* 294-pPG430) 155

FIGURE 4.64. The effect of the number of generations on the plasmid copy number (*E.coli* M5248-pSCC2) 155

FIGURE 4.65. The overall yield factor Yx/s as a function of initial glucose concentration for the recombinant strains *E.coli* 294 and , *E.coli* M5248. 160

FIGURE 4.66. The overall yield factors Yp/x as a function of initial glucose concentration S_0 for the recombinant strains *E.coli* 294 and *E.coli* M5248 160

FIGURE 4.67 The effect of initial glucose concentration on the maxiumum 162

specific growth rate for the *E.coli* 294 strain

FIGURE 4.68. The effect of initial glucose concentration on the maxiumum specific growth rate for the *E.coli* M5248 strain 162

FIGURE 4.69. Structure of the three compartment model for fermentation of *E.coli* 294 165

FIGURE 4.70. Comparison of experimental data and simulation of G-compartment (S_0= 10g/L) 169

FIGURE 4.71. Comparison of experimental data (RNA/60%)and simulation of A compartment (S_0=10g/L) 169

FIGURE A.1. Linear Relation between dry cell weight and OD_{590}nm for *E.coli* M5248 cells 176

FIGURE A.2. Linear Relation between dry cell weight and OD_{595}nm for *E.coli* 294 cells 176

FIGURE B.1. Calibration chart for Bradford assay 177

FIGURE E.1. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 1g/L) 190

FIGURE E.2. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 5g/L) 191

FIGURE E.3. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 10g/L) 191

FIGURE E.4. Curve fitting applied to the biomass formation and 192

the glucose consumption data for *E.coli* 294 (S_0 = 15g/L)

FIGURE E.5. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 30g/L) 192

FIGURE E.6. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 1g/L) 193

FIGURE E.7. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 5g/L) 193

FIGURE E.8. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 10g/L) 194

FIGURE E.9. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 15g/L) 194

FIGURE E.10. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 30g/L) 195

FIGURE E.11. Curve fitting applied to RNA and *Eco*RI formation data for *E.coli* 294 (S_0 = 10g/L) 195

LIST OF TABLES

		Page
TABLE 2.1.	Characteristics of restriction and modification systems	6
TABLE 2.2.	Examples of recognition sequences of class II restriction endonucleases	9
TABLE 2.3.	Conditions altering activity of restriction endonucleases	11
TABLE 2.4.	Amino acid composition of *Eco*RI endonuclease	16
TABLE 2.5.	Culturing requirements and conditions for the growth of microorganisms producing restriction endonucleases	20
TABLE 2.6.	Purification Steps Employed for Some of the Restriction Enzymes	26
TABLE 2.7.	Microbiological process design parameters	35
TABLE 2.8.	Fermentation process design parameters	36
TABLE 2.9	Parameters and other factors in rDNA fermentation	37
TABLE 2.10	Model equations for inhibition free single substrate limiting kinetics	45
TABLE 2.11	Model equations for substrate independent growth kinetics	46

TABLE 2.12 Model equations for inhibition kinetics 47

TABLE 4.1. The effect of the time of induction on the production of *Eco*RI endonuclease from *E.coli* 294 (pPG430) 88

TABLE 4.2. The effect of IPTG concentration on the production of *Eco*RI endonuclease from *E.coli* 294 (pPG430) 88

TABLE 4.3. The effect of temperature-shift on the *Eco*RI enzyme activity 91

TABLE 4.4. Purification of *Eco*RI endonuclease from the recombinant strains 97

TABLE 4.5. Maximum specific growth rates for recombinant and host cells and growth ratios of *E.coli* M5248 112

TABLE 4.6. The maximum specific growth rates (μ_m in hr^{-1}, T-shift from 1-5 hours) 116

TABLE 4.7. The effect of temperature downshift tp 32°C on the *Eco*RI endonuclease activity 117

TABLE 4.8. The maximum specific growth rates of recombinant and host cell cultures in different medium composition 120

TABLE 4.9. The percentage of plasmid containing cells at the late exponential phase in different selective and non-selective media 122

TABLE 4.10. Comparison of the growth ratios and the growth rate differences of recombinant and host cells 123

TABLE 4.11. The effect of different culture media on the *Eco*RI enzyme activity 124

TABLE 4.12. Comparison of experimental and calculated values of the growth rate differences and segregational instability 133

TABLE 4.13. Comparison of the fermentation results of both recombinant strains 136

TABLE 4.14. The overall yield factors Yx/s and Yp/x for the recombinant *E.coli* 294 and *E.coli* M5248 strains 159

TABLE 4.15. Parameter values obtained from the modelling of maximum specific growth rate vs. initial glucose concentration data 163

TABLE A.1. Dry weight data for *E.coli* 294 and *E.coli* M5248 cells 175

TABLE B.1. Standart data for Bradford assay 177

TABLE C.1. Experimental data for biomass and substrate concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 1, 5, 10, g/L 178

TABLE C.2. Experimental data for biomass and substrate concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 15, 20, 30, g/L 179

TABLE C.3. Experimental data for biomass and substrate concentrations for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 1, 5, 10, g/L 180

TABLE C.4. Experimental data for biomass and substrate concentrations for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 15, 30, g/L 181

TABLE D.1. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 1g/L. 182

TABLE D.2. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 5g/L. 183

TABLE D.3. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 10g/L. 184

TABLE D.4. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 15g/L. 185

TABLE D.5. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 200g/L. 186

TABLE D.6. Experimental data for *Eco*RI activity, plasmid DNA, RNA and total protein content concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 30g/L. 187

TABLE D.7. Experimental data for *Eco*RI activity for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 188

1, 5, 10g/L.

TABLE D.7. Experimental data for *Eco*RI activity for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 15, 30 g/L. 189

LIST OF SYMBOLS

A_1	Absorbance without enzyme at 340 nm
A_2	Absorbance with enzyme at 340 nm
$C_{D\text{-glucose}}$	Concentration of glucose (g D-glucose/lt sample)
C_{IPTG}	Concentration of IPTG (mM)
C_N	Plasmid copy number
d	Derivative
F	Dilution factor
k_i	Kinetic parameters (i=1,2...)
K_i	Saturation constant (i=1,2...)
K_p	Product inhibition constant
K_s	Substrate inhibition constant
LB	Liquid broth
m	Number of generations
OD_x	Optical density (x refers to wavelength at with optical density measured)
P_+	Fraction of plasmid-bearing cell concentration over total cell concentration
P_-	Fraction of plasmid-free cell concentration over total cell concentration
P_{-0}	Fraction of plasmid-free cell concentration over total cell concentration at time zero
R	Segregation rate
r_i	Reaction rates (i=1,2...)
S	Substrate concentration (g/L)
S_0	Initial substrate concentration (g/L)
S_{final}	Final substrate concentration (g/L)
S_m	Maximum substrate concentration (g/L)

t	time (hour)
T_{shift}	Temperature shift period (hour)
X	Biomass concentration (g/L)
X_+	Concentration of plasmid-bearing cells (g/L)
X_-	Concentration of plasmid-free cells (g/L)
X_{final}	Final biomass concentration
$Y_{X/S}$	Overall yield factor of biomass concentration over substrate consumption (g cells/g substrate)
$Y_{P/X}$	Overall yield factor of product concentration over substrate consumption (g protein/g cells)
α	Single cell growth rate ratio
ε	Absorbance values
ϕ_P	Fraction of plasmid harboring cells
γ_{ii}	Stoichiometric constants (i=1,2..)
μ	Specific growth rate (hr^{-1})
μ_+	Specific growth rate of plasmid-bearing cells (hr^{-1})
μ_-	Specific growth rate of plasmid-free cells (hr^{-1})
μ^+_{max}	Maximum specific growth rate of plasmid containing cells (hr^{-1})
μ^-_{max}	Maximum specific growth rate of plasmid free cells (hr^{-1})
μ_m	Maximum specific growth rate (hr^{-1})
μ^+_m	Maximum specific growth rate of recombinant cells (hr^{-1})
μ^-_m	Maximum specific growth rate of host cells (hr^{-1})
μ_{mi}	Specific growth rate derived from model i (i=1,2..)
θ	Probability of plasmid loss

1. INTRODUCTION

Restriction enzymes are very important tools of the recombinant DNA technology. Especially, type II restriction endonucleases are very important, since they can cleave DNA at sequence specific sites. A considerable number of highly efficient expression vectors have been engineered in the past few years to obtain high yields of those enzymes encoded by cloned genes of procaryotic origins (Botterman et al., 1985). Strong controllable promoters are preferred in the construction of the expression vectors to enhance the synthesis of the product. The induction conditions are considered to be important parameters in the production of the recombinant proteins and they need to be carefully optimized to increase the yield of the product.

*Eco*RI restriction endonuclease has proven extremely useful as a reagent in the analysis and manipulation of DNA molecules and therefore it is one of the most widely used type II restriction enzymes. In addition to the natural overproducer of *Eco*RI, *E.coli* RY13, genetically modified overproducing strains were also used to produce enzyme (Greene et al., 1974).

In this study, a parametric study of the fermentation by two different genetically modified *Escherichia coli* strains was conducted to optimize the recovery of *Eco*RI restriction endonuclease.

The first strain used was *E.coli* 294 carrying plasmid pPG430, and is an overproducer of the *Eco*RI endonuclease. The induction of the *Eco*RI endonuclease by the addition of IPTG to the medium was optimized. The second strain used was *E.coli M5248* carrying plasmid pSCC2, also an overproducer of the *Eco*RI protein with sufficient genetic stability. *Eco*RI synthesis is induced by transferring the culture to elevated temperatures. The induction parameters were optimized for the synthesis of *Eco*RI endonuclease by *E.coli M5248* (pSCC2). The purification of *Eco*RI endonuclease was accomplished under similar conditions for both strains by using a procedure that was developed by modifying and improving the methods available in the literature (Botterman et al., 1985; Cheng et al.,

1984; Luke and Halford, 1985). Production yields of *Eco*RI endonuclease from both of these overproducing strains were compared. The effect of the various culture conditions on the stability, specific growth rates and *Eco*RI activities of the recombinant plasmids expressing *Eco*RI endonuclease were also studied in both strains.

E.coli 294 (pPG430) strain was chosen for investigating the effect of limiting substrate concentrations on cell growth and on each of the intracellular cell components. Experimental data were obtained on the total protein and total RNA concentrations, plasmid content, *Eco*RI endonuclease activity, substrate consumption, total and viable cell mass. These data were used to develop and verify the mathematical models for recombinant protein production under limiting substrate concentrations. Two different mathematical approaches were used to analyze growth kinetics: unstructured and structured modelling of the biophase. In the unstructured models which were tested for both recombinant *E.coli* strains, the microorganism was regarded as a single reacting species with fixed composition. A three-compartment structured model was developed including the RNA concentration, total protein content and *Eco*RI endonuclease activity along with substrate concentration.

Section 2 gives a literature review of restriction endonucleases in general and of the purification of *Eco*RI endonuclease in particular. The design parameters of recombinant fermentation processes and the kinetics of recombinant cell growth and product formation are also summarized in this section. The experimental methods and the purification procedures used are described in Section 3. The results of experimental studies and modeling are presented and discussed in Section 4. Section 5 gives the major conclusions that can be drawn and the recommendations for future work.

2. THEORETICAL BACKGROUND

2.1 Restriction Endonucleases

2.1.1 General Information on Restriction Endonucleases

Enzymes are an essential feature of natural biological systems. They are complex proteins having the capability to act as a catalyst in the formation or break down of natural compounds. Enzymes that break down nucleic acids are called nucleases. There are two types of nucleases classified according to the way in which they break down a linear nucleic acid molecule. Exonucleases can digest nucleic acid molecules from their ends and endonucleases can break nucleic acid molecules into pieces by cutting within the molecule. The most widely used endonucleases for cutting nucleic acid molecules into pieces are restriction endonucleases (Hale and Mangham, 1991).

Restriction endonucleases are part of the natural defense mechanisms of bacteria against incoming DNA which may be from viruses or plasmids from a foreign population of cells. These enzymes were first recognized by their ability to restrict the growth of certain viruses in certain strains of *E.coli*. The apparent occurrence of these enzymes among some prokaryotes suggests that a search of any prokaryotic family is likely to yield restriction endonucleases (Schildkraut, 1993).

Each organism having a gene which encodes for a restriction endonuclease is presumed to have a cognate DNA modifying enzyme which protects its own DNA from the potentially lethal effects of the restriction endonuclease (Figure 2.1). Modification involves

methylation of certain bases at a very limited number of sequences within DNA which constitute the recognition sequences for the restriction enzyme. During the growth cycle, the host DNA becomes methylated and protected from the restriction system in the presence of the modifying methylase (Heitman, 1993). The recognition sequence is the nucleotide sequence which is required for cleavage, and the site of cleavage is the position of the cleavage with relation to the recognition sequence.

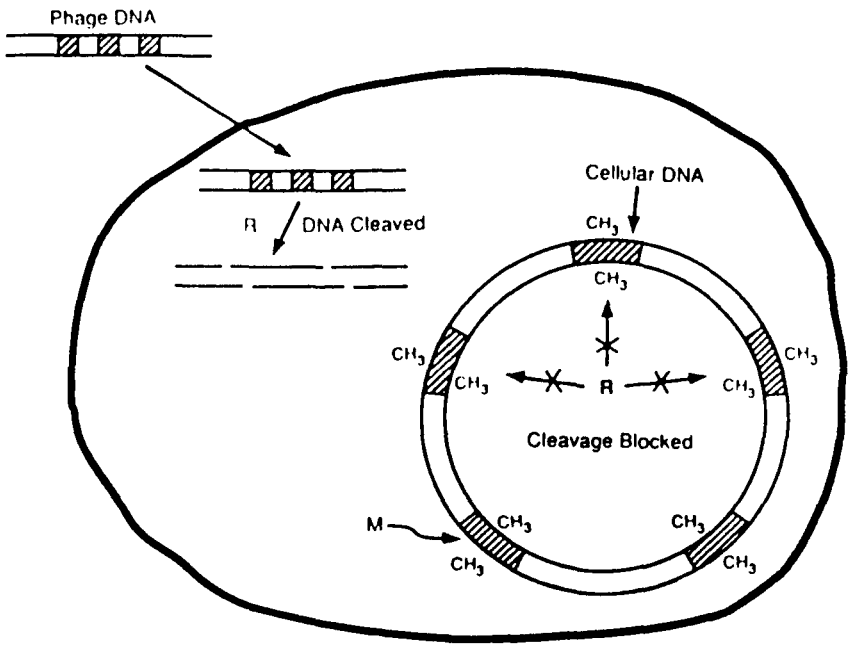


FIGURE 2.1. Restriction-modification. Restriction-modification systems typically consist of an endonuclease (R) and a methyltransferase (M). Hatched boxes represents specific DNA recognition sites modified by methylation or subject to endonuclease cleavage. The host chromosome is protected by methylation whereas unmethylated foreign DNA that enters the cell is restricted (Heitmen, 1993).

Nomenclature for the restriction enzymes follows a simple convention. The names start with three letters in which the first letter of the genus is followed by the first two letters of the species of the source cells. They are followed by an indication of the strain and a roman number indicating which one of the enzymes from that strain the name refers to (Bailey and Ollis, 1986). For example, the restriction enzyme *Eco*RI refers to the first activity isolated from strain R of *E.coli*.

Several enzymes purified from different microorganisms appear to recognize the same nucleotide sequence. These enzymes are called **isochizomers**. Although the recognition sequences of the isochizomers are the same, they may vary with respect to the site of cleavage (*Xma*I and *Sma*I), sensitivity to methylation (*Msp*I and *Hpa*II) and in cleavage rates at various sites (*Xho*I and *Pae*R7I) (Berger, 1987). If the isochizomers are also the same with respect to the site of cleavage, in each group of the isochizomers only one of these enzymes is chosen as a prototype, often because it was the first to be isolated, or in some cases because its purification scheme leads to a better yield than the one obtained with the others.

Based on the properties of the purified enzymes, restriction-modification systems fall into three classes: type I, type II and type III, and their main characteristics are summarized in Table 2.1.

Type I enzymes are the most complex ones and they consist of three subunits each having restriction, modification and specificity. They form a pentameric multifunctional enzyme ($R_2 M_2 S$) and require complex cofactors as Mg^{+2}, ATP, S-adenosylmethionine for DNA cleavage. Their cleavage are at random sites distant from the site of recognition. These enzymes form a loop between sites of recognition and cleavage by unwinding and lose their activity after a single round of DNA scission (Heitman, 1993).

Type III enzymes consist of two subunits having restriction and modification functions. They form a methyltransferase of modification subunit alone and a methyltransferase endonuclease complex containing both restricition and modification subunits. Although Mg^{+2} and ATP are required for DNA cleavage as in type I enzymes, S-adenosylmethionine stimulates enzymatic activity without being an absolute requirement.

They cleave the DNA at unique sites but yield partial digests that do not proceed to completion. Finally, type III enzymes modify only one DNA strand (Heitman, 1993).

TABLE 2.1. Characteristics of restriction and modification systems (Howe,1995)

	Type I	**Type II**	**Type III**
Composition	Multienzyme complex with endonuclease, methylase and specificty subunits.	Separate enzymes: endonuclease is a homodimer, methylase is a monomer	Methylase subunit provides specificity and it is as heterodimers with endonuclease subunit. Function as methylase-endonuclease
Cofactors	Mg^{2+} , ATP, SAM (needed for cleavage and methylation)	Mg^{2+} (for cleavage) SAM (for methylation)	Mg^{2+} , ATP (for cleavage), SAM (for methylation)
Recognition Sites	Asymmetric	Symmetric	Asymmetric
Cleavage	Variable distance from recognition site	Within recognition site	Precise distance from recognition site
Number of systems	Several, grouped into a few families	Hundreds	Few

SAM: S- Adenosyl methionine

Type II systems are much simpler. These endonucleases contain only one type of subunit. Type I and type III enzymes cut the DNA molecule distant from their recognition sites. Only type II enzymes recognize and cleave the DNA molecule at a specific sequence. Type II restriction endonucleases recognize a particular target sequence in a duplex DNA molecule and break the polynucleotide chains within that sequence to give rise to discrete DNA fragments of defined length and sequence (Primrose,1991). They are homodimers

distinct from and without having an identity to their methyltransferases. Each enzyme requires a single cofactor: Mg^{+2} for cleavage and S-adenosylmethionine (SAM) for methylation.

One of the common type II enzymes is *Eco*RI restriction and modification enzymes which are separable proteins specified by different genes. They have simple catalytic requirements; in vitro modifications require only unmodified DNA and S-adenosylmethionine (SAM) whereas restriction of unmodified DNA requires only Mg^{+2}. The enzyme forms a stable sequence-specific complex with DNA containing the recognition site. The methylase is composed of a single polypeptide chain and only the monomer has been detected in solution (Rubin and Modrich, 1977; Needles et al., 1989).

The restriction endonucleases often make a staggered cut to leave molecules that have single-stranded ends. These are called **sticky** or **cohesive ends**. They enable any two DNA fragments to be joined, because they are complementary to each other. Some of the type II enzymes cut both strands at the same place resulting in **blunt-ended** fragments. Sticky ends can form a relatively stable structure with respect to blunt-ends and are desirable on the DNA molecules to be ligated together in a gene cloning experiment. As a result of the cleavage by *Eco*RI enzymes, sticky ends are produced as shown in Figure2.2.

The restriction endonucleases have been characterized primarily with respect to their recognition sequences and cleavage specificity rather than their protein properties because of their practical uses for molecular dissection of DNA. Most of the restriction endonucleases recognize sequences that are four-six nucleotides in length, but there are also many of them with seven-eight base recognition sites. The recognition sites for a number of type II enzymes given by Roberts and Macelis (1991) are arranged in tabular form in Table 2.2. Unique recognition sequence is not the only characteristic of a restriction endonuclease that should be considered when evaluating its usefulness. The symmetry in their recognition sequence is also an important characteristics of these enzymes. The recognition sites of most of them contain an axis of symmetry and in most cases all the bases within the site are uniquely specified. The symmetrical recognition sequence of these endonucleases have been termed **palindromes.** There are two types of palindromic sequences, perfect and imperfect. A perfect palindromic sequence is a simple, invariable palindrome. An imperfect

palindromic sequence contains variation at particular locations within the recognition sequence. For example, the imperfect palindrome GT(A/C)(G/T)AC represents four permutations, two of which are perfect palindromes (GTATAC, GTCGAC) and two of which are non-palindromes (GTAGAC, GTCTAC) (Schildkraut,1993).

There are many factors affecting the cleavage of the DNA by the restriction endonucleases: methylation, secondary structure of the substrate and the buffer composition are the most important ones.

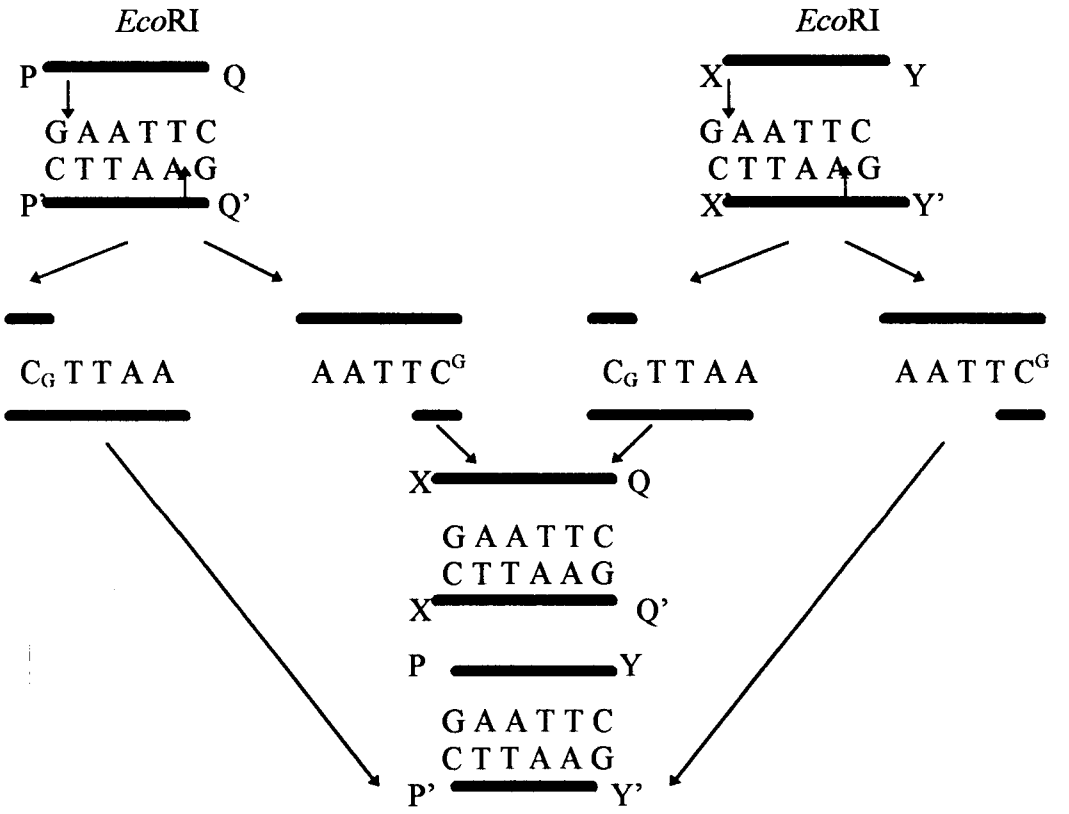


FIGURE 2.2. *Eco*RI restriction enzyme's sticky ended cut (Modrich and Zabel, 1976)

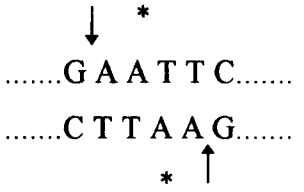
TABLE 2.2. Examples of recognition sequences of class II restriction endonucleases

Enzymes	Recognition sequence	Origin	Isochizomers
***Apa*I**	GGGCC'C C'CCGGG	*Acetobacter pasteurianus*	*Bsp*120I
***BamH*I**	G'GATCC CCTAG'G	*Bacillus amyloliquefaciens*	*Bst*I
***Dra*I**	TTT'AAA AAA'TTT	*Deinococcus radiophilus*	*Aha* III
***Hpa*II**	C'CGG GGC'C	*Haemophilus parainfluenzae*	Hap II, Bse II, *Fin* II
***Sal*I**	G'TCGAC CAGCT'G	*Streptomyces albus*	*Rhp* I, *Xam* I
***Sph*I**	GCATG'C C'GTACG	*Streptomyces phaeochromogenes*	*Spa* I, *Bbu* I, *Pae* I
***Xba*I**	T'CTAGA AGATC'T	*Xanthomonas badrii*	
***Bgl*II**	A'GATCT TCTAG'A	*Bacillus globigii*	
***Dpn*I**	GA'TC CT'AG	*Diplococcus pneumoniae*	Nan II
***Eco*RI**	G'AATTC CTTAA'G	*Escherichia coli RY13*	*Rsr* I, *Sso* I
***Eco*RII**	'CCGG GGCC'	*Escherichia coli R245*	*Aor* II, *Atu* II, *Sgr* II
***Hind*III**	A'AGCTT TTCGA'A	*Haemophilus influenza*	*Bbr* I, *Bpe* I
***Mae*III**	'GTNAC CANTG'	*Methanococcus aelicus*	
***Pst*I**	CTGCA'G GA'CGTC	*Providencia stuartii*	*Xma* II, *Xph* I, *Xor* I
***Pvu*II**	CAG'CTG GTC'GAC	*Proteus vulgaris*	*Bav* I
***Sau*3AI**	'GATC CTAG'	*Staphylococcus aureus*	*Mpo* I
***Taq*I**	T'CGA AGC'T	*Thermus aquaticus*	*Tfl* I
***Xma*I**	C'CCGGG GGGCC'C	*Xanthomonas malvacearum*	*Sma* I

Methylation of bases in DNA are the basis of the modification activity in the restriction-modification system. Methylated bases commonly encountered include N^6-methyladenine, 5-methylcytosine, 5-hydroxymethylcytosine and N^4-methylcytosine. In

general prokaryotic and some simple eukaryotic organisms contain both adenine and cystosine methylated DNA; higher multicellular eukaryotes such as mice and man are only known to contain cytosine methylation, and some organisms, such as *S.cerevisiae* and *D.melanogaster*, lack detectable DNA methylation. (Heitman,1993). Thus DNA methylation is not essential for viability. In case of cells expressing a sequence-specific endonuclease, it is required to protect the host chromosome from destruction. Since restriction endonucleases do not cut generally where specific bases within their recognition sites are metyhlated (Kaszubska et al., 1989), it is the basis of the protection mechanism in the cell to save its own DNA. Methylation at other positions do not affect the cleavage (Primrose,1991).

For example, *Eco*RI resriction and modification enzymes recognize the double stranded six nucleotide sequence and cleave each strand between the G and A residues as marked below.



Restriction is a consequence of double strand cleavage within this sequence shown by arrows, while the modification is the result of methylation of the 2 adenine residues adjacent to the axis of symmetry shown by asterisks.

The secondary structures in the DNA may also interfere with recognition by the restriction endonuclease. It may result in less efficient cleavage of some of the recognition sites than the others within the same molecule

The specificity of many restriction endonucleases is also affected by the buffer composition used. Several restriction endonucleases have been found to cleave nucleotide sequences which are similar but not identical to their recognition sequence depending upon the environmental conditions. These enzymes exhibit **star activity** which is the activity resulting from reduced specificity under suboptimal conditions. High endonuclease

concentrations, substitution of manganese for magnesium, low ionic strength, high pH or presence of organic solvents can cause to observe the star fragments rather than the fragments of normal recognition site. For the same enzyme, there have been different recognition sites under different incubation conditions (Table 2.3).

TABLE 2.3. Conditions altering activity of restriction endonucleases (Perbal,1988)

Enzymes	Conditions
***Ava*I**	Glycerol, enzyme concentration
***Bam*HI**	Glycerol, enzyme concentration, low salt concentration, Mn^{2+} substituted for Mg^{2+}
***Bst*I**	Glycerol, enzyme concentration
***Eco*RI**	Glycerol, enzyme concentration, low salt concentration, Mn^{2+} substituted for Mg^{2+}
***Eco*RV**	Glycerol
***Hha*I**	Glycerol, enzyme concentration
***Hind*III**	Low salt concentration, Mn^{2+} substituted for Mg^{2+}
***Hpa*I**	Glycerol, enzyme concentration
***Pst*I**	Glycerol, enzyme concentration
***Pvu*II**	Glycerol, enzyme concentration
***Xba*I**	Glycerol, enzyme concentration

Star activity was first observed and best studied with *Eco*RI endonuclease (Heitman,1992). The recognition specificity of *Eco*RI endonuclease is reduced to the tetranucleotide sequence as

-N A A T T N-
-N T T A A N-

when ionic strength is lowered and pH adjusted to 8.5. It was also reported that under high glycerol concentrations, *Eco*RI star activity takes place at G G A T T T, A A A T T T, G A A T T T and G A A T T A sites. Another interesting site for star activity of *Eco*RI is that it

could be any hexanucleotide sequence that contains no more than one base substitution from the normal recognition site, providing that the substitution does not result in an A to T or T to A change in the central portion of the recognition site (Mayer, 1978). Buffer compositions are very important also because it effects the yield of the enzyme. *Eco*RI enzymes have a tendency for the formation of an insoluble aggregate of the protein inside cell and precipitate out of the solution which often leads to a decrease in the yield of the enzyme. The enhanced solubility of the enzyme at high NaCl concentrations shows that the solubilization requires the distruption of ionic interactions between protein molecules (Terry et al.,1983; Luke and Halford, 1987).

2.1.2 Structure and Function of Type II Restriction Enzymes

In studying the structure and function of type II restriction enzymes, the focus is on the nature of the specific and non-specific protein-DNA interactions, the details of the catalytic mechanism, and the physical nature of the coupling between recognition and catalysis. Nearly 2000 type II endonucleases have been identified from a wide variety of bacterial genera.

There have been some speculations on restriction endonucleases as they may fulfill a purpose other than restriction. One suggestion has been that DNA scission by restriction enzymes may stimulate recombination *in vivo* similar to the ways in which they have been used as *in vitro*. DNA recombination allows microorganisms to reassort genes contributed to progeny. Recombination also permits the cell to repair DNA lesions, such as breaks and gaps, which destroy the genetic information borne by a chromosome. Several sequence-specific endonucleases generate double-straned breaks in DNA that promote recombination and mobilization of introns, retrotransposons. One of the other views on the evolutionary origin for restriction endonucleases and methyltransferases is that they arose independently, and subsequent events including gene transfer and mutagenesis paired and then tailored endonuclease and methyltransferase specificities (Heitman, 1993).

The structure of the enzyme-DNA complex of the restriction endonucleases are determined by X-ray crystallographic analysis. The detailed models for mechanisms of DNA recognition and catalysis are provided by X-ray crystallography. The X-ray structures of type II restriction endonucleases-DNA substrate complexes have been studied in the absence of the essential Mg^{2+} as their only cofactor to prevent DNA scission. Enzyme-product structures have been studied when Mg^{2+} has been diffused into the crystals for cleavage to occur. General knowledge of DNA-protein interactions has been provided by different authors (Heitman, 1993; Winkler, 1992; Rosenberg, 1991).

*Eco*RI was the first restriction endonuclease to be isolated in quantity, and consequently the first for which the structural information became available. The structure of the gene encoding *Eco*RI and biochemical characteristics of this enzyme have been extensively studied. These studies made the structure-function analyses possible (Rosenberg, 1991).

It has been shown that the endonuclease acts as a dimer of identical subunits. Each monomer in *Eco*RI has a primary β-sheet with α-helices on both its sites. Four of the five strands of the β-sheet are parallel; however the location of the single antiparallel strand makes it possible to divide the sheet into topological motifs that have identifiable functional roles.The first three strands , β1, β2, β3 are antiparallel and form a β-meander that contains the cleavage site. Strands β3, β4, β5 form the basis of a nucleotide binding fold that interacts with the base pairs. The stereo drawings of the Cα atoms of *Eco*RI endonuclease and the substrate DNA are shown in Figure 2.3 (Winkler,1992; Rosenberg, 1991).

The palindromic nature of the target sequence reflects the two fold rotational symmetry of the dimeric protein. If the target sequence is modified by the methylation, the sequence is then resistant to endonuclease. The resistance of the half-methylated site protects the bacterial host's own duplex DNA from attack immediately after semiconservative replication of the fully methylated site until the modification methylase can once again restore the daughter duplexes to the fully methylated site.

a

b

FIGURE 2.3. Stereo drawings of Cα atoms of *Eco*RI endonuclease and the substrate DNA, a and b shows two different site view of the complex. *Eco*RI coordinates were taken from entry 1RIE in the Brookhaven Protein Data Bank.

The interaction of the *Eco*RI endonuclease with the recognition site comprises hydrogen bonds and a set of Van der Waals contacts. The hydrogen bonds can be divided

into three classes: hydrogen bonds between bases and aminoacid side chains; hydrogen bonds between bases and amide groups in the main polypeptide chain; water-mediated hydrogen bonds in which a water molecule is sandwiched between a base and amino acid side chains in a sequence specific manner. The Van der Waals contacts are observed between thymine and methyl groups (Jeltsch et al., 1992).

One of the important observations reported for *Eco*RI endonuclease is that it shows two distinct modes of action. The first one is the isosteric mode which is encountered when *Eco*RI binds sites containing base analogs or some base mispairs, that is the *Eco*RI-canonical site complex. The second one is the adaptive mode which is observed when *Eco*RI binds *Eco*RI* sites, which differs substantially from the pattern obtained for canonical complex. The features of the adaptive patterns are these: first, all *Eco*RI* sites produce very similar changes; second, the changes involve both the disapperance of some points of strong interference and apperance of new ones; and finally the adaptive pattern is highly asymmetric, whereas the canonical pattern reflects the symmetry of the *Eco*RI site. This feature suggest that there are two basic recognition-site conformations for *Eco*RI, canonical-isosteric and adaptive (Winkler,1992; Rosenberg, 1991)

Although *Eco*RI endonuclease and methylase recognize a common hexanucleotide sequence, the physical and catalytic properties are quite distinct. The native molecular weight of *Eco*RI methylase is reported as Mr of 39,000 and it is a monomer. The endonuclease is composed of a subunit with Mr of 28,500 and exists as dimers and tetramers in solutions. Although both forms can give rise to endonuclease activity, they are not equally catalytically active. Under dilute conditions employed for catalytic assays, the active form of the endonuclease is most likely the dimer (Modrich and Zabel, 1976).

The two enzymes have also different isoelectric points. In contrast to basic character of the methylase, the endonuclease is an acidic protein with an isoelectric point of 6.3. The mechanism of methylation indicates that methyl transfer occurs to only one DNA strand of the *Eco*RI sequence prior to enzyme dissociation (Rubin and Modrich,1977).

The amino acid composition of the homogeneous *Eco*RI endonuclease is shown in Table 2.4 (Modrich and Zabel, 1976).

TABLE 2.4. Amino acid composition of *Eco*RI endonuclease

Amino acid	Moles per Mr=28,500	Amino acid	Moles per Mr=28,500
Alanine	15.8	**Lysine**	19.6
Arginine	12.2	**Methionine**	4.5
Aspartic	36	**Phenylalanine**	9.93
Half-cystine	1.5	**Proline**	5.84
Glutamic	25.3	**Serine**	19.5
Glycine	20.2	**Threonine**	9.4
Histidine	4.9	**Trytophan**	1.86
Isoleucine	19.1	**Tyrosine**	8.21
Leucine	24.8	**Valine**	16.6

2.1.3. Purification of Restriction Enzymes

Purification of the restriction enzymes is necessary to be able to use them as a reagent in research or in analytical diagnostic works. On the other hand, the purification of an enzyme may not always be the end point in itself, but it can be the mean to obtain a pure enzyme for further studies. These studies may be on the enzymology such as on characterization or on its structure-function relationships, or structure determination of the enzyme. The amino acid sequence and the three dimensional structure of the protein can be determined through the use of X-ray crystallography, protein sequence and computer modeling.

The enzymology studies, considering the activity of the enzyme, will require relatively small amounts of active protein, but elimination of any interfering activities will be essential. It will be necessary to have the enzymes as pure as possible for enzyme specificity in particular. **Structural studies** will require larger amounts of highly pure

enzyme. For **structural and functional analysis**, not only highly pure but also highly active enzyme is necessary. For many **mapping, cloning and recombinant DNA experiments** restriction enzymes need not to be homogeneously pure. They must be free of other restriction enzymes, exonucleases and endonucleases as well as inhibitors of enzyme e.g. kinases, ligases, phosphatases.

The economy of purification of enzymes highly depends on the achievable product concentration. The amount of purified enzyme will not only depend on the amount of the starting material but will also depend on yield. The final yield of enzyme depends to a great extent upon the number of fractionation steps employed during purification procedure. Several fractionating steps will lead to a great loss in the final yield. However by minimizing the number of steps, the final purity and the concentration of the protein will be lower. Thus by judicious choice of the individual steps, yield can be maximized with a minimal loss in purity.

The fractionation steps during purification will depend on the source and the cellular location of the products that are located either within the cell or are secreted from the cell. For intracellular products, conditions of growth affect the disruption kinetics. The stage at which cells are harvested is important since the cells harvested during the log phase are more susceptible to homogenization than those in the late stationary phase.

A general purification scheme may be proposed for restriction endonucleases such as preparation of the clarified cell extract after disrupting the cell wall following harvesting of cells, then removal of DNA either by precipitation or by ion-exchange chromatography and finally fractionating the proteins by use of different chromatographies such as ion exchange, adsorption or affinity (Greene et al., 1974).

2.1.3.1. Sources and Growth Conditions. There are two main sources for restriction endonucleases. The first one is the natural hosts and the second one is the genetically engineered microorganisms (Howe,1985).

Bacteria remain the main natural sources of the restriction endonucleases, since the restriction endonuclease and modification methytransferase enzymes are part of the natural

defense system to destroy the foreign DNA entering the cell. A few of the DNA-site specific enzymes have also been reported in eukaryotes like *Saccharomyces cerevisiae, Pichia membranefaciens, Spirulina platensis* and *Chlamydomonas reinhardtii*. Even a restriction endonuclease, HsaI, has been reported from human beings (Dubey et al., 1987). The yield of these enzymes from the natural hosts are very low and it requires that purification be started with a large amount of cell paste such as 0.1-1kg.

The use of screening techniques allowed the search for suitable producing strains for restriction enzymes and to improve the strains genetically through mutations. On the other hand, the efficient development of suitable vector systems has led to the generation of the genetically modified microorganisms producing specific restriction enzymes in large amounts.

The ability to purify the desired enzyme depends largely on the ability to grow pure cultures of the microorganism selected. The selected microorganism is cultivated in suitable medium at appropriate environmental conditions of pH, temperature, oxygen demand, appropriate time period. Cultural conditions have a significant effect on physiology and metabolism on the microorganisms and their synthesizing capacity.

Cell growth is related to three processes: uptake of nutrients into the cell through the cell wall (substrate diffusion), conversion and utilization of nutrients (reactions), and release of products. The growth cycle of a microbial populations cultured in batch conditions consists of four phases: lag phase, exponential growth phase, stationary phase and the death phase. The length of the lag phase depends on the nutrient composition of the medium, on the size and the age of the inoculum. When the growth begins, the cells multiply rapidly resulting in an exponential increase in the cell number. As the amount of rate limiting substrate reaches a minimum, the cell mass approaches a maximum level. It is during the stationary phase that biomass reaches its maximum concentration; some cells divide while others die. Eventually, due to accumulation of toxic product and nutrient depletion, the population can not sustain itself and the death phase begins (Bailey and Ollis, 1986).

In order to maximize production of a particular protein the optimization of culture conditions is essential. The culture media should contain all the materials necessary for the

microorganism to move quickly into the exponential phase of growth. The requirements and conditions of growth vary widely for different microorganisms which produce restriction enzymes and examples are shown in Table2.5. In general the activity of the restriction enzymes are at their maximum before the onset of the stationary phase (Dubey et al., 1987). In some cases the amount of enzyme per cell does not seem to vary substantially during the growth cycle, and it is therefore possible to allow the cells to enter into stationary phase before harvesting. The choice of cultivation time and the conditions of the harvesting of the cells on a large scale are essential for high recovery of restriction enzymes.

2.1.3.2. Preparation of the Cell Extract. The cell wall is a thick, rigid coat formed outside the cell membrane and is the major barrier to the release of any intracellular proteins. The cell membrane of both gram-positive and gram-negative bacteria gives shape and some protection to the cell and also acts as a regulatory filter for the transport of materials into and out of the cell.

Disruption of the cell wall may be achieved by chemical, physical or mechanical means in order to release intracellular proteins prior to purification. Small quantities of cells can be disrupted by means of physical methods such as simultaneous homogenization in blenders, agitation with abrasives, use of liquid or solid extrusion forcing either the cell suspension or the cell paste at high pressure through a narrow orifice. Excessive foaming can cause denaturation of protein or the addition of abrasives may cause important problems during the protein purification, limiting the use of these methods. The cells can also be disrupting by exposing them to extreme temperature conditions or reacting them with different chemicals. Desiccation, explosive decompression, osmotic shock, freeze thaw, alkali treatment or treatment with detergent and solvents can also be used to destroy the cell wall. These methods are not preferred because of the cost or the protein denaturation and instability. Lytic enzyme treatment can also be employed in cell disruption. Lysozyme for example which catalyses the hydrolysis of the glycosidic bonds in peptidoglycan can be used in bacteria. The disadvantage is that it is very expensive to be used in large scale (Broda, 1979).

TABLE 2.5. Culturing requirements and conditions for the growth of microorganisms producing restriction endonucleases (Dubey et al., 1987).

Restriction Enzyme	Organism	Culture Media (composition/liter)	Cultivation Conditions
***Aac*I, *Aae*I**	*Acetobacter liquefaciens*	Glucose: 5g Glycerol: 15g Peptone: 5g Yeast Extract: 1g Malt extract: 1g	Temperature: 30°C pH: 6.8 Aeration: Continuous agititation
***Bam*NI, *Bam*HI**	*Bacillus amyloliquefaciens*	Bactopeptone: 8g Yeast extract: 5g $KH_2PO_4.3H_2O$: 3g Sucrose: 5g	Temperature: 37°C pH: 7.0 Aeration:10-50 ml air/min Harvesting:After middle or end of log phase
***Eco*RI**	*Escherichia coli RY13*	Tryptone: 10g Yeast extract: 5g NaCl: 10g Glucose: 5g K-phosphate: 5mM	Temperature: 37°C pH: 7.0 Aeration: Maximum aeration Harvesting: Early stationary phase
***Pst*I**	*Providencia stuartii*	Tryptone: 15g Yeast Extract: 10g NaCl: 5g	Temperature: 37°C pH: 7.0 Harvesting:After 16 hour Inoculum age and concentration: 2% overnigth culture
***Taq* I**	*Thermus aquaticus YT1*	Castenholz TYE Medium	Temperature: 70°C pH: 7.4 Harvesting: Stationary phase

The sonication method is the most widely used technique. A high transient pressure is produced when the suspensions are subjected to ultrasonic vibrations. When frequencies of 20kHz and above are applied to solutions, they cause areas of rarefaction and compression which rapidly interchange. As the gas bubbles collapse, shock waves are formed. The protein release is almost proportional to the acoustic power and independent of

cell concentration. At very high cell concentrations insufficient mixing may lower disruption. Cell breakage occurs at an exponential rate, dependent on exposure time. As the contact time increases, heat is produced. In order to minimize the effect of heat, the cell suspension should be kept on ice and circulated away from the ultrasonic probe to the outside of the vessel where it is cooled. When the cells are being ruptured by ultrasonication technique, the temperature must not be permitted to go above 10°C. Figure 2.4 summarizes the steps required for the preparation of crude extracts from microorganisms.

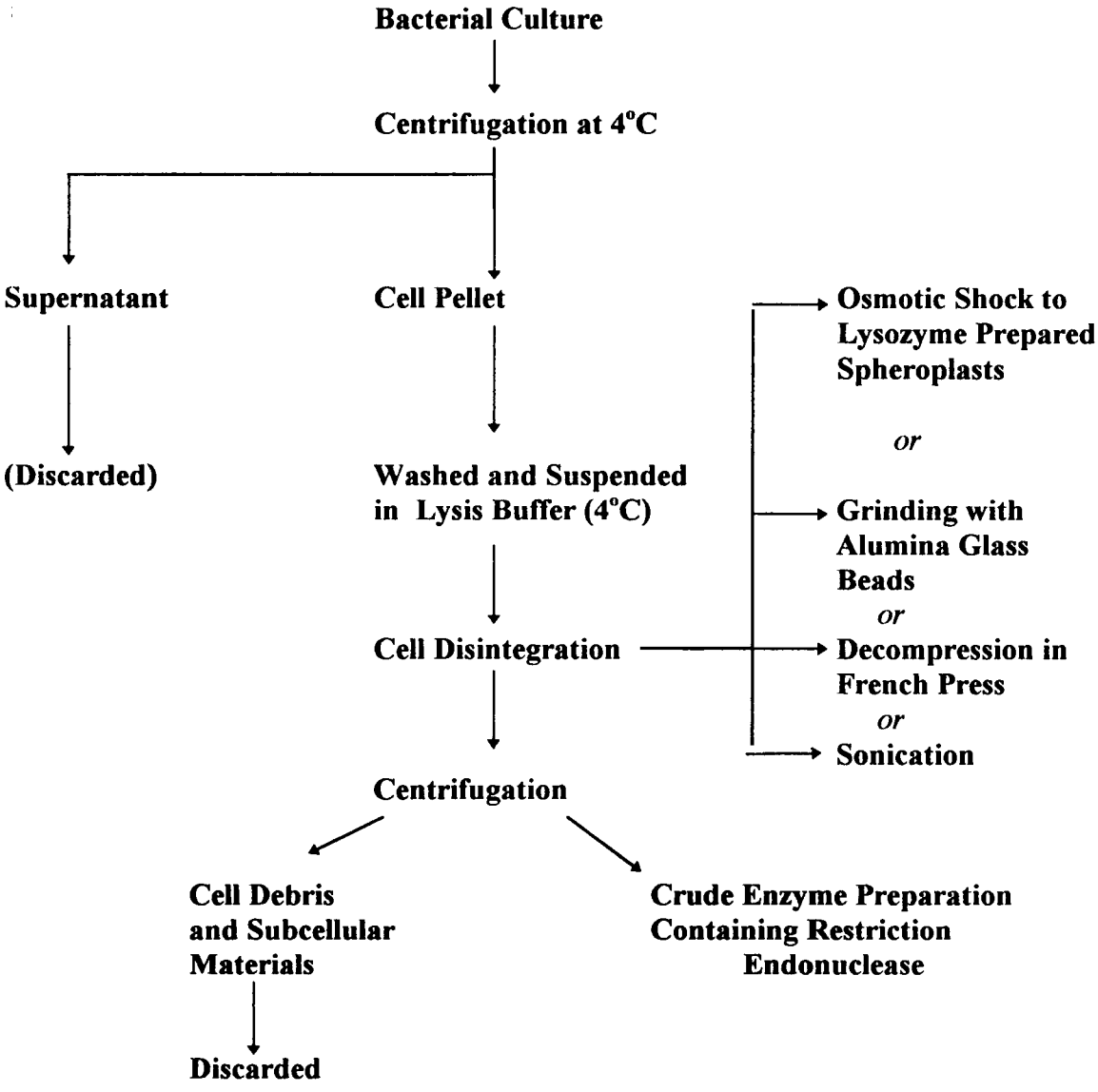


FIGURE 2.4. The steps involved in the crude extract preparation for purification of intracellular restriction endonuclease from microorganisms (Dubey et al., 1987)

The cell breakage is usually monitored in a number of ways depending on the inherent properties of the cells. The loss of turbidity can be measured by measuring the optical density at 550nm. It is appropriate for *Haemophilus*, *Bacillus*, *Neisseria* and many other bacterial generations which do not aggregate or clump. The release of proteins can be followed by suspending cells, centrifuging the debris, and measuring the optical density of the supernatant at 260-280nm. The degree of cell lysis can also be monitored by phase contrast microscopy and this method is found to be the most suitable technique to follow the distruption of blue-green algae cells (Schildkraut, 1991).

Once the cells have been opened, it is necessary to remove unbroken cells and cell debris. Cell extract can be clarified by centrifugation and the supernatant containing the desired enzyme can be processed to purify. The storage of the crude extract is not recommended. When storage is needed at this stage, crude extracts can be stored at -20°C or 4°C in the presence of 50% glycerin and/or in the presence of a proteolytic inhibitor such as PMSF (20μg/ml) (Schildkraut, 1991).

Since the presence of nucleic acids changes the chromatographic properties of many enzymes, it is essential to remove as much of the nucleic acids as possible from the cell extracts before the further purification of the restriction endonucleases.

Several methods have been used to remove nucleic acids from the environment. Gel filtration chromatography at high ionic strength has been widely used. Precipitation is another way of removing nucleic acids. Streptomycin sulfate and polyethyleneimine are common agents used for the preferential precipitation of nucleic acids. Polyethyleneimine itself interferes with chromatography; therefore, it needs to be removed from the solution by ammonium sulfate precipitation (Flaschel and Friehs, 1991).

2.1.3.3. Chromatographic Techniques. The selection of a particular form of chromatography to achieve a separation is dependent on the material to be isolated, and often several chromatographic methods may be used sequentially to achieve the complete purification of a compound. Basically, all chromatographic systems consist of two phases. One is the stationary phase which may be solid, gel, liquid or a solid-liquid mixture which is

immobilized. The second phase which is mobile may be liquid or gaseous and flows over or through the stationary phase (Wilson and Goulding,1986).

The **major types of chromatography** are adsorption chromatography, partition chromatography, gas-liquid chromatography, ion-exchange chromatography, exclusion chromatography and affinity chromatography.

There are three **major modes of chromatographic separations**; column chromatography, thin layer chromatography and paper chromatography. In column chromatography, the stationary phase is packed into glass or metal columns whereas in thin layer chromatography, the stationary phase is thinly coated onto glass or plastic plates. In paper chromatography the stationary phase is supported by the cellulose fibers of the paper sheet. Most of the chromatographic separations in biochemical applications are carried out using the column mode.

There are **important factors** that have to be considered **when using the column chromatography technique** for separation. It is necessary to maintain the temperature of the column constant during the experiment. Packing of columns is one of the critical factors in achieving a successful separation. It is normally carried out by gently pouring the slurry of the stationary phase into a column and letting the packing settle evenly. It is necessary to maintain a layer of solvent above the column surface to prevent "running dry". The sample is very carefully applied by using a pipette, capillary tubing, a syringe or a peristaltic pump and allowed to flow into the column without disturbing the column material. Overloading the column with sample should be avoided in order to prevent irregular separation. It is difficult to generalize about the ideal column height to diameter ratio and the total bed volume. They both influence the amount of material which can be separated on the column and in practice will need to be determined by systematic trial and error. During the elution process, it is essential that the mobile phase flows through the column at a stable rate. The separation of the product can be achieved through the elution of a single solvent (isocratic separation) or gradient elution in which the pH, ionic concentration or the polarity of the eluant is continuously changed (Wilson and Goulding,1992).

In ion exchange chromatography, the principle feature is the attraction between oppositely charged particles. Ion exchange separations are mainly carried out in columns packed with an ion-exchanger. There are two types of ion-exchanger, namely cation and anion exchangers. Cation exchangers possess negatively charged groups and these will attract positively charged molecules. Anion exchangers have positively charged groups which will attract negatively charged molecules. First, the solution containing the protein to be purified is given to the column and then the bound ion is displaced by the protein via a substitution reaction. Finally an appropriate buffer with a higher ionic concentration is used to release the protein from the column. The common ion-exchangers are synthetically prepared derivatives of cellulose such as diethylaminoethylcellulose, carboxymethylcellulose and phosphocellulose.

Gel filtration column chromatography separates protein molecules according to their molecular weight and/or size. The pore size is such that the larger protein molecule pass around the beads and elute from the column rapidly. Smaller protein molecules enter the resin pores, and therefore move more slowly through the column. Commonly used gels include dextrans (Sephadex), agarose (Sepharose, Bio-Gel A, Sagavac) and polystyrenes (BioBeads S).

Affinity chromatography purification exploits the unique property of extremely specific biological interactions to achieve separation and purification. The technique requires that the material to be isolated is capable of reversible binding to a specific ligand which is attached to an insoluble matrix. Under correct experimental conditions when a complex mixture containing the specific compound to be purified is added to the insolubilised ligand, only that component will bind to the ligand. All other components can therefore be washed away and the compound subsequently recovered by displacement from the ligand.

In adsorption chromatography, the adsorbent used is a solid that has the property of holding molecules at its surface. It has a very large surface area in relation to particle size. Adsorption can fairly be specific so that one solute may be adsorbed selectively from a mixture. Separation of components by this method depends upon differences in both their degree of adsorption by the adsorbent and their solubility in the solvent used for separation.

The most widely used and effective adsorbent for protein purification is a form of crystalline calcium phosphate, hydroxyapatite.

High pressure liquid chromatography is a recently developed technique which may employ the principles of adsorption, partition, ion exchange exclusion and affinity chromatography. This technique has the advantages of high efficiency and fast speed of resolution, and it is used in the column chromatography mode.

For most restriction endonucleases, purification is obtained by one or more steps of column chromatography. One of the common column materials is used phosphocellulose which has the advantages of selectivity for nucleic acid binding proteins and high capacity. Heparine covalently linked to agarose has also been used as an alternative to phosphocellulose. Proteins which bind to nucleic acids apparently recognize heparin as a nucleic acid analog and bind to the heparin-agarose column because of this affinity rather than by ion exchange. DEAE-cellulose may also be used for the same purpose under well defined conditions (Aiken and Gumport, 1988).

Currently employed purification steps for purification of some of the restriction enzymes are summarized in Table 2.6 (Greene et al., 1978; Pirrotta and Bickle, 1980; Jack et al., 1991)

Depending on the enzyme and on the degree of purity required, it may be necessary to employ a second chromatographic step. DNA-cellulose or agarose and hydrophobic interaction chromatography have also been employed for restriction enzyme purifications. The last system might be particularly useful for those enzymes that have a tendency to precipitate in low salt concentrations, since hydrophobic matrices are loaded at relatively high salt concentrations.

2.1.3.4. Enzyme Assays. Several assays have been used to determine the activity of the enzymes and also to identify the fractions which contain the enzyme. Ideally, the enzyme assay should be simple, highly specific and rapid, allowing many fractions to be screened for activity prior to the next stage of the purification. Many enzymes can rapidly be assayed

spectrophotometrically but problems of interference may be encountered with other components in impure samples.

TABLE 2.6. Purification Steps Employed for Some of the Restriction Enzymes

Restriction Enzyme	Organism	Separation Steps
***Bam*HI**	*Bacillus amyloliquefaciens H*	1.Crude extract preparation 2. Phosphocellulose - salt elution 3.Hydroxyapatite chromatography 5.Heparin-sepharose salt elution
***Eco*RI**	*Escherichia coli RY13*	1. Crude extract preparation 2. PEI eluate 3. Phosphocellulose
***Pst* I**	*Providencia Stuartii*	1. Crude extract preparation 2. Nucleic acid precipitation with PEI 3. NH_4SO_4 fractionation 4. DEAE-cellulose
***Taq* I**	*Thermus aquaticus YTI*	1. Crude extract preparation 2. Phosphocellulose 3. Heparin-sepharose 4. Hydroxyapatite

The use of radioisotope techniques have been involved in the determination of the activity of the type II restriction enzymes. Duplex DNA in a nicked or cleaved form was trapped as a single stranded DNA on a nitrocellulose filter after being heated in the assay which was used by Center et al (Dubey et al, 1987). The radioactivity is counted after the filter is dried. This method takes several hours to complete the assay. DeFilippes (Dubey et

al, 1987) has used the same principle except that the double-stranded DNA is adsorbed on a small amount of hydroxyapatite instead of being trapped on a filter paper.

The infection DNA assay was used by Takanami (Dubey et al., 1987). The number of infection centres scored by transfection of infective bacteriophage DNA is proportional to the amount of phage DNA used under appropriate conditions. Endonucleolytic cleavage of in tact DNA molecule destroys the plaque forming ability. This transsection method is therefore applicable to assay of restriction endonucleases as well as other endonucleases. It has been used for preparation of enzymes from *Haemophilus* strains.

The conversion of duplex circular DNA to linear DNA in the presence of ethidium bromide or polynucleotide kinase were also used as fluorimetric assay procedures for type II restriction enzymes (Dubey et al., 1987).

The most common assay procedure for restriction enzymes is electrophoresis of digests of DNA on slab gels of polyacrylamide or on agarose.

The amount of enzyme present in a particular preparation is conventionally expressed not in terms of units of mass or moles but in terms of units based upon the reaction that the enzyme promotes. The *international unit* (U) is defined as the amount of enzyme required to completely digest a standard amount (usually 1 microgram) of a standard type of DNA (often bacteriophage lambda, or a specific plasmid) in a given time (usually one hour) under optimal assay conditions (Wilson and Goulding, 1992).

Kelly et al (Dubey et al., 1987) have used the definiton in which one unit of activity corresponds to 1.0 pmol of sites cleaved per hour at 37°C for HhaII. Smith et al. (Dubey et al., 1987) defined a unit of BamHI endonuclease as the amount of enzyme needed to digest 1.0 pmol of phosphodiester bonds per minute in superhelical pMB29 DNA at 37°C. Takanami et al. (Dubey et al., 1987) defined one unit of enzyme as the activity which digested 0.01 unit of DNA within 30 minutes at 37°C.

Digestion is complete when the total length of all the resulting fragments is equal to the full size of the undigested DNA. Total lengths greater than the undigested DNA are

characteristic of incomplete digestion. Digesting DNA molecules containing many sites may require more units of enzyme than the amount required to digest the same mass of a DNA containing fewer sites.

The protein assays combined with enzyme assays provide information on the specific activity of the purified protein. Specific activity is expressed as total units of enzyme per total amount of protein.

2.1.3.5. Criteria of Purity. Purification of a restriction endonuclease can be carried out to different degrees depending upon the ultimate use. For mapping or comparative studies, an enzyme need only be purified to the point where digests of DNA give sharp bands on agarose or polyacrylamide gels. For DNA sequencing work, this is not enough; enzymes have to be completely free from contaminating exonucleases, single-stranded-specific nucleases, etc.

The absence of contaminant exo- and endonucleases capable of digesting DNA either at random or at specific sites appropriate to the purified enzyme are usually tested by incubating DNA under different reaction conditions. Overdigestion and the cut-ligate-recut tests are widely used as quality control tests by the manufacturers.

In the overdigestion test, the enzyme preparations are tested by incubating the substrate DNA with an excess amount of enzyme for longer periods than the reaction time. In the cut-ligate-recut test; the quality of the preparations are determined by what fraction of the products can be religated and whether the religated molecules can still be cut with the enzyme.

The purity of the final enzyme solution are determined under the denatured conditions by subjecting to SDS-PAGE polyacrylamide electrophoresis. The apparance of single and sharp protein bands on the gel, is always desired at the end of the purification (Pirrotta and Bickle, 1980)

2.1.3.6. Stability and Storage of Restriction Endonucleases. It is observed that complete digestion of high molecular weight cellular DNA requires more than the theoretical amount

of restriction enzyme according to the standard unit definition. At this point, stability of the enzyme becomes an important parameter over extended periods of incubation time. *Bam*HI, *Kpn*I, *Pst*I and *Sma*I were among the least stable enzymes under extended periods of incubation time (Schilkraut, 1991).

Several restriction enzymes have been found to be stable around the neutral pH. During fractionation they can easily lose their activity at temperatures higher than 10°C. Purified restriction enzymes are stored at -20°C in buffers containing 50% glycerol. Highly purified enzyme preparations can lose activity upon storage. Addition of gelatin or bovine serum albumin to a final concentration of 50-100 μg/ml will prevent the loss of activity. Some of the enzymes such as *Eco*RI and *Pst*I also require a neutral detergent as 0.2% Triton X-100 to stabilize them (Pirrotta and Bickle, 1980).

2.2. Purification of *Eco*RI Endonuclease

*Eco*RI endonuclease is either purified from its natural source and its overproducing derivatives or from genetically modified strains constructed to express the enzyme at high yield by using recombinant DNA techonology.

An early construct used as a source of *Eco*RI restriction endonuclease and modification methylase is the *E.coli RY13* strain. It is a derivative of a widely used *E.coli 1100 (K12)* strain. The strain was made by r_B^+ m_B^+ and a derepressed fi^+ R factor carrying the *Eco*RI restriction and modification genes introduced by conjugation and selection for the appropriate drug resistance genes. The R factor in *E.coli RY13* carries drug resistance genes for sulfathiazole and streptomycin. The productivity of the enzyme is generally very low at the end of the different purification procedures applied to this strain. Several investigators have used *E.coli* RY13 strain to produce *Eco*RI restriction endonuclease ending with different specific activities and the yields depend upon the purification protocol which they have used (Greene et al., 1974; Modrich and Zabel, 1976; Sümegi et al., 1977;

Bingham et al., 1977; Rubin and Modrich, 1980; Vlakatis and Bouriotis, 1991; Mehra et al., 1993).

High expression of *Eco*RI enzymes were achieved by using different constructions and different host organisms resulting in a overproducing system of the enzyme. Botterman and Zabeau (1985) have constructed an overproducing strain of *E.coli* that can synthesize about 30% of its cellular protein as *Eco*RI restriction endonuclease compared to about 0.01% of the total protein in the natural isolates of *E.coli*. Cheng et al (1984) also worked on a different construction of another plasmid which resulted in a 1.5% synthesis of its cellular protein as *Eco*RI. Although the endonuclease was under the control of the p_L promoter in each case, possible reasons for this 20 fold difference could include the following: the distance between the p_L promoter and the cloned gene for *Eco*RI were 600bp longer on the plasmid of Cheng et al. (1984), cloned genes are on different plasmids and different host strains in each case resulting in a different plasmid copy number. Despite a 20 fold lower synthesis of *Eco*RI in the construction of Cheng et al (1984), *Eco*RI expression in the strain of Botterman and Zabeau leads to an inactive and insoluble protein in large amounts Their yield was only 4 fold higher than the yield obtained from the strain of Cheng et al (1984).

Botterman and Zabeau (1985) have constructed the *Eco*RI endonuclease overproducing strain of *E.coli* in three steps. First, they have fused the N-terminal part of the *Eco*RI endonuclease gene to lacZ gene fusion vector. Secondly the p_L promoter was randomly positioned upstream of this hybrid gene to optimize the level of expression. Finally high level synthesis of the product was achieved by linking different fragments containing the p_L promoter and the N-terminal part of the *Eco*RI endonuclease gene to the corresponding C-terminal part leading to the construction of suitable host strains in which the plasmids were stable.

In the construction of Cheng et al. (1984), a 2210bp *Bam*HI fragment containing *Eco*RI endonuclease and methylase gene was isolated. The fragment was ligated to plasmid pKC30 downstream from the bacteriophage λ p_L promoter and then transformed into *E.coli N99/*λ^+ (Shimatake and Rosenberg, 1981). The orientation of the insert was determined by digestion of recombinant plasmids with *Bgl*II restriction endonuclease. A

plasmid containing *Eco*RI structural genes in correct orientation for expression is named as plasmid pSCC2. Plasmid pSCC2 was introduced into strain M5248, a defective λcI857 lysogen deleted for prophage sequences rightward from cro which fails to produce phage structural proteins upon thermal inactivation of the cI repressor. The overproduction of *Eco*RI endonuclease and methylase, upon thermal induction was reported (Figure 4.2).

Another overproducing construction of *Eco*RI was investigated by Wu and coworkers (1993). They have employed the combination of three different types of multicopy plasmids for the overproduction of *Eco*RI in *E.coli JM103* strain.Plasmid pRK248cI carries the gene for temperature sensitive cI857 repressor protein and tetracyclin resistant gene. Plasmid pEcoR4 encodes the DNA methylase of *Eco*RI and chloramphenicol resistance. Plasmid pRIF309 carries the gene for *Eco*RI under control of p_L promoter.

In the present study, two different overproducing strains *E.coli*, carrying two different *Eco*RI endonuclease expression vector systems were investigated in terms of their productivity under the same circumstances. The first one is the construction of Cheng et.al. (1984), that is, *E.coli* M5248 cells carrying plasmid pSCC2, and the second one is the construction of Boyer et.al. (1981), that is, *E.coli* 294 cells carrying plasmid pPG430. In pPG430 plasmid, *Eco*RI structural genes are placed under the control of lacUV5 promoter. It is 4633 bp long and a derivative of pBR322 plasmid carrying the β-lactamase origin. Plasmid pMB1 contains the endonuclease and the methylase genes.

The restriction endonuclease *Eco*RI has been purified in many laboratories using different purification procedures. The protocols are summarized below in chronological order.

Greene et al (1974) have developed a method for the purification of *Eco*RI endonuclease from the *E.coli* RY 13 strain. They have used the sonication method for cell distruption. Following the addition of streptomycin sulfate and then ammonium sulfate to the sonicated cell suspension, phosphocellulose chromatography was used. Their method involves chromatographic separations on hydroxyapatite, DEAE-cellulose columns followed by a second hydroxyapatite and finally Sephadex columns. Their yield was 13 U/gcell.

Modrich et al (1976) have increased the yield of the enzyme to 190 U/gcell from *E.coli* RY 13 strain by modifying the method of Greene et al (1974). They have used streptomycin sulfate, ammonium sulfate fractionation and then chromatographic separation on phosphocellulose, hydroxyapatite and DNA cellulose.

Sümegi et al (1977) have developed a procedure to purify restriction endonuclease *Eco*RI by using *E.coli* RY 13 cells with a yield of $4.3x10^5$ U/gcell. Their method involves sonication of cells at low ionic strength, precipitation of the endonuclease with Polymin P, elution of the enzyme from the Polymin P precipitate, ammonium sulfate precipitation and chromatography on phosphocellulose.

Bingham et al (1977) have worked with the same strain and reached to a yield of $5.2x10^5$ U/gcell by modifying the above method. After suspending the precipitate of Polymin P in a buffer, they have concentrated the suspension by the use of an ultrafiltration membrane and applied phosphocellulose chromatography.

Cheng et al (1984) have constructed on overproducing strain, *E.coli* M5248, to purify *Eco*RI restriction and modification enzymes. Their method involves streptomycin sulfate and ammonium sulfate fractionations followed by chromatography on phosphocellulose and hydroxyapatite. They have reached a yield of $2.1x10^5$ U/gcell .

Botterman et al (1985) have also used a genetically modified strain of *E.coli*, *E.coli* 1100, to overproduce the *Eco*RI enzymes. They have applied the supernatant of the sonicated cell suspension to phosphocellulose and then to hydroxyapatite chromatography. Their yield was $3.8x10^7$ U/gcell.

Luke and Halford (1985) have worked with the same strain as Botterman et al (1985) and increased the specific activity of the *Eco*RI enzyme four fold by modifiying the procedure above, obtaining a similar yield. They have included a dialysis step which prevents the aggregation of the protein, before the application of phosphocellulose chromatography.

Vlatakis and Bouriotis (1990) have purified *Eco*RI endonuclease again from the natural strain, *E.coli* RY 13. They have applied the sonicated cell suspension to a DNA-cellulose column and then onto an oligonucleotide-ligand affinity chromatography. They have ended up with a yield of $1.8x10^5$ U/gcell.

Mehra et al (1993) have applied the dye-ligand chromatography. They have used Cibacron Blue-F3GA-agarose column followed by a hydroxyapatite column to purify *Eco*RI endonuclease from *E.coli* RY 13. Their yield was $3x10^4$ U/gcell.

2.3 Recombinant Fermentation Processes

One of the aims of modern biotechnology is the overproduction of intracellular and/or extracellular products by using recombinant microorganisms.

Recent developments and advances in the area of genetic engineering and biotechnology have shown that there are unlimited possibilities for their potential applications to various gene products by means of: (1) transferring genes to bacterial host organisms from other gene sources; (2) improving gene expression efficiencies; (3) amplifying gene copy numbers; and (4) improving purification efficiency (Ryu and Lee, 1987).

2.3.1 Bioprocess Design Parameters

Both genetic and environmental factors affect cell growth and product formation (Ryu and Lee, 1987).The genetic factors that render unique characteristics to a given

recombinant strain include DNA replication, transcription and translation. The environmental factors include temperature, aeration, agitation, pressure, apparent viscosity, osmotic pressure, pH, carbon and nitrogen sources, medium composition, growth factors, precursors, inducers, metabolites with toxic and /or inhibitory effects, etc. These parameters are also highly interactive and their relationship to overall productivity are illustrated in Figure 2.5.

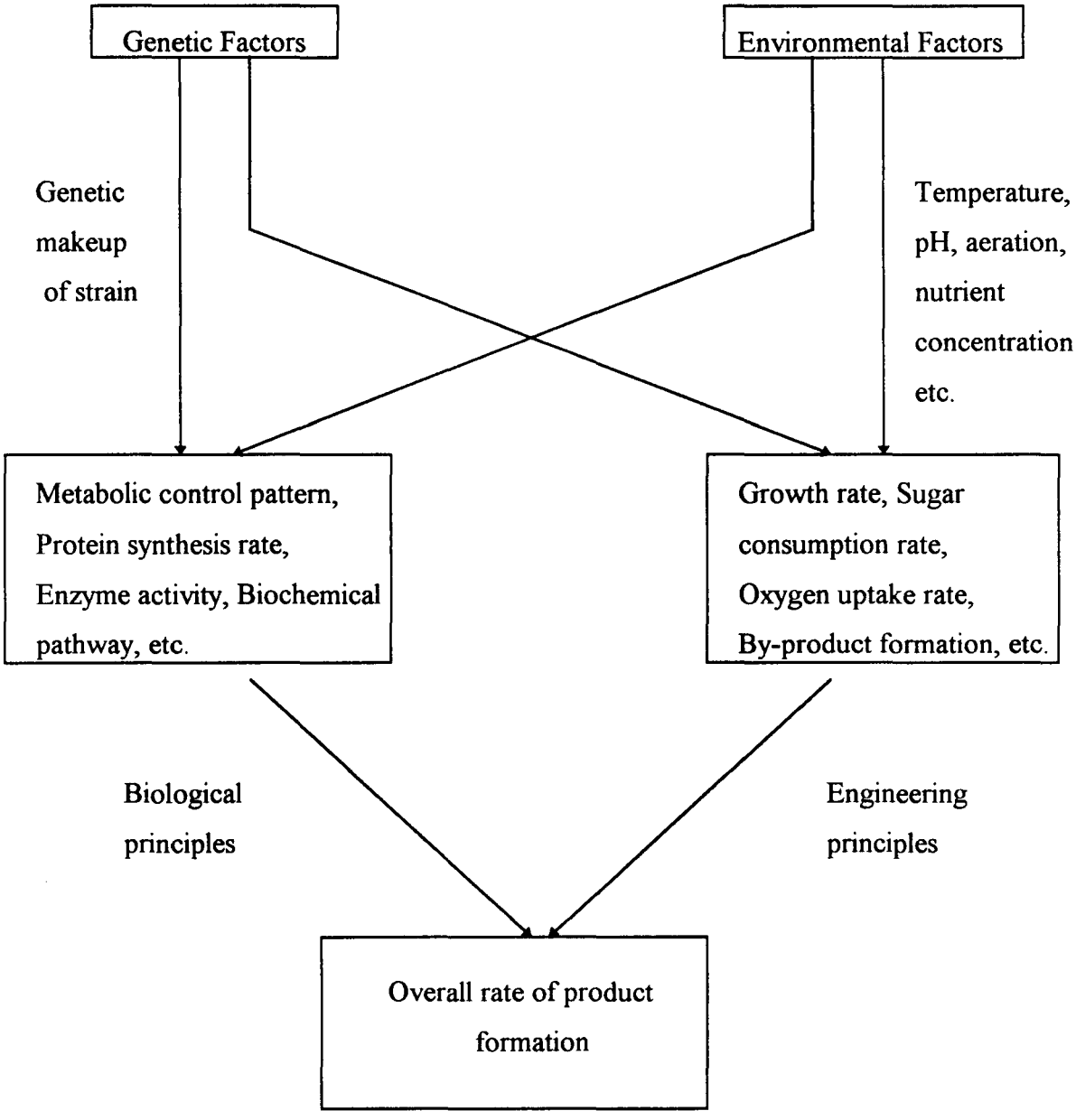


FIGURE 2.5. Schematic diagram of the influence of genetic and environmental factors on overall product synthesis (Ryu and Lee, 1987)

The selection list of important microbiological process design parameters that deserve consideration in optimizing recombinant fermentation processes are presented in Table 2.7

TABLE 2.7. Microbiological process design parameters (Ryu and Lee, 1987)

1	Gene Dosage
	* Plasmid copy number
	* Regulation of replication
2	Transcription Efficency
	* Promoter strength
	* Regulation of transcription
3	Translation Efficiency
	* Nucleotide sequence of ribosome binding site
	(Shine-Dalgarno sequence)
	* Distance between Shine-Dalgarno sequence and ATG or GTG initiation codon
	*Secondary structure of mRNA
	* Codon Usage
4	Stability of recombinant DNA
	* Stability of cloned gene
	* Stability of plasmid
5	Stability of mRNA
	* Structure of mRNA
	*Nuclease
6	Stability of Protein
	* Structure of protein or gene product
	* Protease
7	Host cell
	* Host / vector interaction
	* Metabolic activity of host cell
8	Others
	* Protein secretion efficiency
	* Proper termination between genes

In recombinant fermentation processes, basically the same variables that are used in the traditional fermentation processes should be considered. The important fermentation process parameters are listed in Tables 2.8. and 2.9.

TABLE 2.8. Fermentation process design parameters (Ryu and Lee, 1987)

1	**Medium design**
	* Concentration of carbon sources
	* Concentration of nitrogen sources
	* Concentration of trace elements
	* Other nutrients and growth factors
	* Precursors and /or effector concentration
2	**Bioreactor design**
	* Agitator speed
	* Air flow rate
	* Viscosity of culture fluid
	* Power input
	* Medium feed rate
	* Cooling water flow rate and heat transfer
3	**Other culture conditions**
	* pH
	* Temperature
	*Dissolved oxygen (DO) level
	* Dilution rate

TABLE 2.9. Parameters and other factors in rDNA fermentation (Ryu and Lee, 1987)

1	**Inoculum preparation**
	*Stable maintenance of recombinant cells during inoculum development
2	**Media design**
	* Medium design for maintenance of recombinant cells
3	**Oxygen transfer**
	* Increased oxygen demand & sensitivity to oxygen perturbation
4	**Process control and optimization**
	* Fine control process variables (pH, temperature, DO level, etc) * Optimization of induction time
5	**Others**
	* Reduced shear rate when protein is secreted into medium * Optimization of dilution rate in continuos fermentation

2.3.2. Optimization of Recombinant Fermentation Processes

Both the microbiological process variables and fermentation process variables affect productivity and must be carefully studied and evaluated when the recombinant fermentation process is to be scaled up and optimized.

In optimizing microbiological process variables, the optimal set of conditions with respect to DNA replication, transcription and translation will have be determined. From an engineering point of view, on the other hand, very little known about scaling up and

optimizing recombinant fermentation processes, and substantial work needs to be done for commercializing new gene products (Ryu and Lee, 1987).

2.3.2.1. Temperature. Temperature optimization is still an important area of investigation for the recombinant fermentation process development. While the optimal growth temperature for the host organism, *E.coli*, is 37°C, the optimal temperature for production of both interferon and insulin was found to be about 30°C (Backman and Ptashne, 1978; Meyer et al., 1984; Emerick et al., 1984). In the case of insulin production, productivity was increased threefold when the cultivation temperature was maintained at 30°C as compared with the productivity obtained when the fermentation temperature was 37°C. This increase was attributable to a significant reduction in product degradation at lower temperature.

2.3.2.2. Time of Induction. The selection of the best induction time, the period of induction and the concentration of the inducer for gene expression during growth cycle are important in optimizing the recombinant fermentation processes.

Early inductions will lead to a waste of energy during protein synthesis, degradation of the protein can easily occur as a result of the accumulation of the target protein and may also be harmful to the host cells. Late inductions will also cause a decrease in the yield and consequently in the activity of the enzyme especially because of the decrease in the number of viable cells.

In the literature, there are a number of studies on the influence of the induction time on gene expression. The work of Neubauer et.al (1992) is on the maximization of the recombinant gene expression in *E.coli* by manipulation of induction time using lactose as inducer. They have found that the time of addition of lactose is very important for the quality of induction, and they have achieved a high yield by adding lactose to the system during the transition from the exponential to the stationary phase. In another study, it was shown that an early induction of the gene expression during the logarithmic phase of growth for production of *Eco*RI endonuclease resulted in about five fold increase in productivity as compared to that of late induction during the stationary phase (Botterman et al., 1985).

2.3.2.3. Oxygen Demand and Dilution Rate. A sufficient oxygen supply to the recombinant cells, especially during the production phase immediately following gene expression, are important in optimizing the recombinant fermentation process.

In a chemostat culture, the dilution rate corresponds to the specific growth rate and in principle these parameters should have the same effect on productivity. However, one needs to distinguish these parameters for plasmid-harboring and plasmid-free cell populations when highly unstable recombinant organisms are used and there is a mixed population with different growth characteristics.

2.3.3. Stability of Recombinant Plasmids

Despite many attempts to optimize recombinant fermentation processes, the instability of recombinant plasmids remains one of the most important problems in the commercial application of many recombinants. There is both the structural instability and segregational instability in rDNA. Structural instability is often caused by deletion, insertion, recombination, or other events, while segregational instability is caused by uneven distribution of plasmids during cell division (Primrose and Ehrlich, 1981). Both types of instability result in a loss of productivity. Thus, it is also important to examine: (1) which parameters determine the degree of instability, (2) which environmental factors affect the instability of recombinants , and (3) how one can make best use of unstable recombinants.

Plasmid stability is defined as the ability of transformed cells to maintain their plasmid unchanged during their growth. In the scale up of fermentation processes using recombinant cells, the main objective is to maintain the plasmid content of the cells unchanged over long periods of time, since it is well known that increase in the level of expression of recombinant genes will cause a decrease in cell growth rates and may result in morphological changes.

The productivity of a bioreactor employing recombinant strains is largely affected by the degree to which the plasmid-free cells are generated and propagated. The plasmid-free cells are generated from plasmid-harboring cells by segregational instability which is caused by defective partitioning of the plasmids between the daughter cells during cell division. Another source of instability originates from the changes in the plasmid itself. The resulting cells either with absent or structurally altered plasmids are non-productive. The plasmid-harboring cells usually grow more slowly than the plasmid-free cells because plasmid-harboring cells have to synthesize more DNA, mRNA and protein. Thus once generated in the fermentor, plasmid-free cells may propagate rapidly leading to a mixed population with plasmid-free cell proportion of the population increasing with each generation.

Although continuous systems are highly productive for many microbial processes, their application is more limited for recombinant organisms because of plasmid instability. Most recombinants proteins are therefore produced by either batch or fed-batch processes (Kumar et al., 1991).

Various strategies have been proposed to improve plasmid stability. Some of these include maintaining selection for antibiotic resistance by the use of antibiotics in the growth medium, complementation of host auxotrophy by incorporating auxotrophic markers on plasmid vectors, lysogenic phage repression and incorporation of suicide proteins whose synthesis is repressed in the presence of the plasmid. Application of whole-cell immobilization to recombinant organisms was also found to improve the stability of plasmids. The use of two-stage cultivation and recycling between different dilution rates with different substrate concentrations have been among the reported methods (Kumar et al., 1991).

The degree of instability can be determined by the growth ratio parameter and the plasmid-loss or segregation rate. The growth ratio, α, is defined as the ratio of the specific growth rate of plasmid-harboring cells, μ^+, to that of plasmid-free cells, μ^-,i.e. μ^+/μ^- . In general, the growth ratio α, is less than one (Ryu and Lee, 1987). The segregation rate parameter, θ, is defined as the ratio of the specific rate of generation of plasmid-free cells from the plasmid-harboring cells (Θ; plasmid loss rate) to the specific growth rate of

plasmid-harboring cells (μ^+). The segregation rate parameter (Θ/μ^+) depends largely on the genetic characteristics.

DuPoet et al (1987) have worked with *E.coli* JM105 cells carrying plasmid pKK223-200. They have concluded that the plasmid copy number and consequently the plasmid stability depends very much on the growth medium used. In minimal medium, a large increase in the copy number with a constant level of enzyme activity was attained compared to a lower copy number with higher enzymatic activity in LB medium.

Iniesta (1988) has tested the effect of oxygen tension on the plasmid stability of plasmid pTG201. Plasmid pTG201 derived from pBR322, has been incorporated into *E.coli* K12 cells. The plasmid had higher stability in cultures supplied with 100% oxygen than in 21% oxygen.

Similarly, Huang and coworkers (1990) have examined the effect of oxygen transport on the plasmid stability in *E.coli* B cells with pTG201 plasmid. They have increased the oxygen transport by increasing the rate of agitation which in turn resulted in higher plasmid stability with increased level of enzyme productivity.

Huang and his coworkers (1993) have also worked on the effect of inducer concentration on plasmid stability. They found that the percentage of plasmid bearing cells dramatically drops following the IPTG induction in *E.coli* DH5a cells containing plasmid pMJR1750. At the end of the four hour induction period, they have obtained a high level of enzymatic activity.

Lavastida et al (1993) have investigated the plasmid instability problem in recombinant *E.coli* JM103 cells and encountered segregational instability. The segregational instability was found to be caused by the accumulation of mutations in the old recombinant strains, and plasmids themselves were not responsible for the instability. Therefore, they have concluded that, in many cases, periodic retransformation of recombinant plasmids into fresh host cells may be a useful method for maintaining plasmid stability.

Lamotte et al (1994) have studied plasmid stability in recombinant *E.coli* HB101 cells producing glyceraldehyde 3-phosphate dehydrogenase (GAPDH). An unusal phenomenon was noticed concerning the plasmid stability of this strain growing in batch culture. A transitory decrease in the percentage of the resistant cells has been noted during the exponential phase of growth which corresponds to a momentary plasmid instability probably due to a gap between the growth rate of the cell and the duplication rate of the plasmid. They have also concluded that the cells harboring a high plasmid copy number are more stable than the cells having a low plasmid copy number in batch cultures.

The above results reported in the literature show that each recombinant system will have its own characteristics and should be carefully investigated to achieve an optimum fermentation process with optimum product yields.

2.4. Modeling of Bacterial Cell Growth and Product Formation

A number of mathematical models have been developed to describe growth and product formation for recombinant microorganism. Models can be divided into two main subgroups: concerning the cellular representations according to the number of components and concerning the cellular properties as being homogeneous or heterogeneous. Cellular representations which are multicomponent are called structured, and single component representations are designated as unstructured. Consideration of discrete, heterogeneous cells constitutes a segregated viewpoint, while an unsegregated perspective considers the average cellular properties. As indicated in Figure 2.6., the actual situation is a structured and segregated one. The most idealized case is the unstructured and unsegregated one (Bailey and Ollis, 1986).

Unstructured models are the simplest. They take the cell mass as a uniform quantity without internal dynamics whose reaction rate depends only upon the conditions in the liquid phase of the reactor. Therefore, the models only contain kinetics of growth,

substrate uptake and product formation. Models which take into account the changes in the internal state of the population and of the effects of this on the growth rate are called **structured models**. The models can be structured on the basis of biomass components such as concentrations of metabolites, enzymes or RNA or by population-related variables describing different morphological types of cells or cell aging. Models with a structure on the population level are also called **segregated models**.

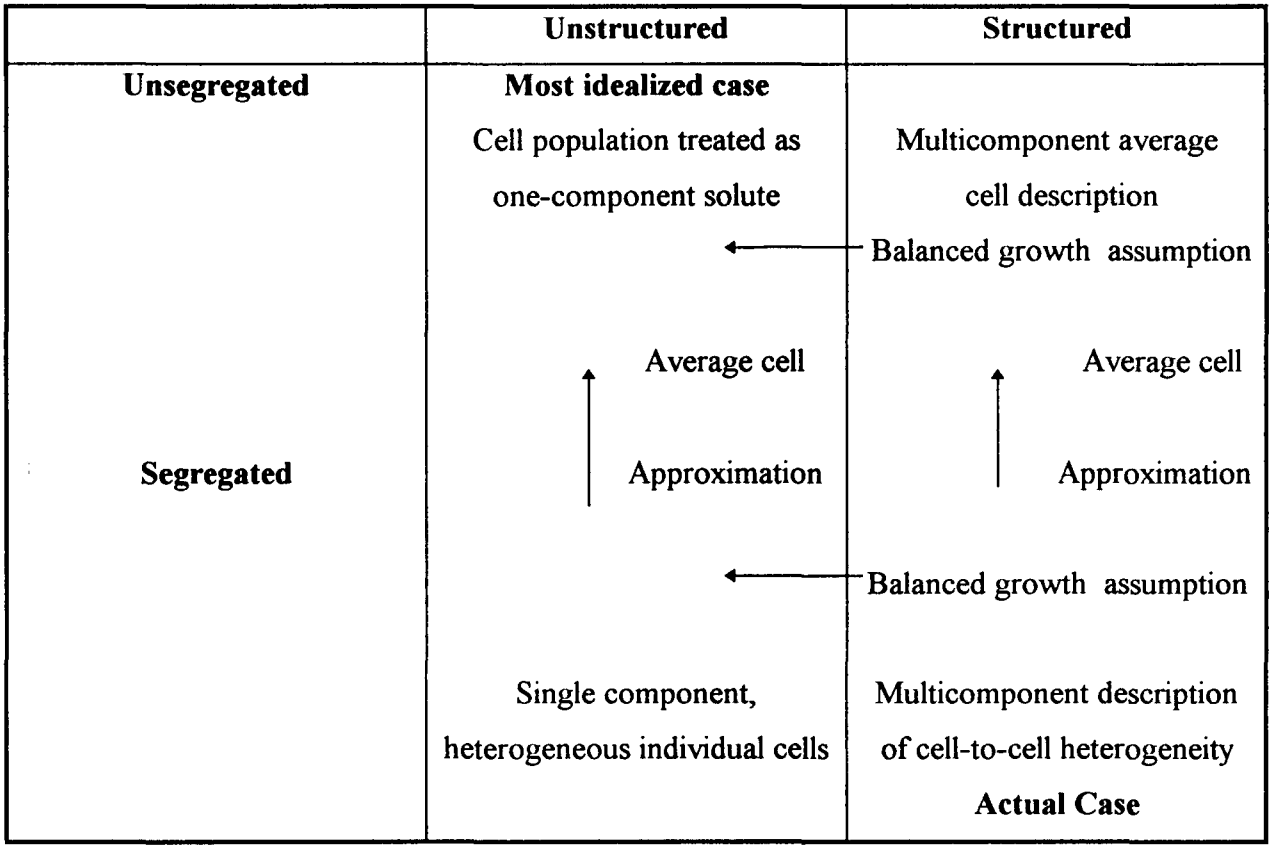


FIGURE 2.6. Different perspectives for cell population kinetic representations

2.4.1. Unstructured Models

In unstructured models, the biological reaction depends directly and solely on macroscopic variables that describe the conditions in the fermentor. The only biotic state

variable is the cell mass concentration.The relatively unspecific modeling of the biotic phase in unstructured models limits their descriptive potency. Only the average growth behavior of the culture is described in the sense of an averaging over the population of the cells and a time averaging over the cell division cycle. In many cases this simplification is reasonable because it is impossible to have exact knowledge of the heterogeneous composition of the biomass and the state of the intracellular systems (Han and Levenspiel, 1988; Lefebvre et al., 1994).

2.4.1.1. Inhibition-Free Single-Substrate Limiting Kinetics. During growth, new cell mass is formed autocatalytically from substrate with specific growth rate, μ:

$$r_x = \mu X \qquad (2.1)$$

where is the net rate of cell mass growth. Several basic types of kinetics for the specific growth rate have been proposed, as summarized in Table 2.10.

Although the equations in Table 2.10 should be taken as formal kinetics for the globally observed behavior of the culture, one can give an interpretation of the special form of certain kinetics. A functional relationship between the specific growth rate μ and concentration of an essential compound was proposed by Monod. The equation is analogous to the standart rate equation for enzyme-catalyzed reactions with single substrate (Michaelis-Menten enzyme kinetics). In the Monod equation, μ_{max} is the maximum growth rate achievable when $s>>K_s$ and the concentrations of all other essential nutrients are unchanged. K_s is that value of the limiting nutrient concentration at which the specific growth rate is half its maximum value. Its applicability to biotechnological processes can be referred to growth rate limiting systems for the substrate. In Monod kinetics, μ approaches its asymptote too slowly to be a proper approximation of experimental data even in simple cases. Teissier and Blackman are proposed related forms of specific growth rate dependence which give better fits to experimental data . The Moser and Vavilin equations are similar to the Monod kinetics with the exception that the reaction for the substrate is not of first order. The Vavilin equation, is often used application in processes with toxic substrates, especially for aerobic treatment of waste water. The equation of Contois is also similiar to Monod kinetics and contains an apparent Michaelis constant which is proportional to

biomass concentration X. Equation of Mason and Milles account for additional diffusion-driven flux of substrate into cell. Powell considered different functions and taked into account substrate diffusion and cell size by a factor K_D.

TABLE 2.10. Model functions for inhibition-free single-substrate limiting kinetics (Rehm et al., 1991)

Model Name, Year Proposed	Kinetic Model	Equation Number
Blackman, 1905	$\left(\frac{\mu}{\mu_{max}}\right) = 0.5\ S/K$ for S< 2K $\left(\frac{\mu}{\mu_{max}}\right) = 1$ for S>2K	2.2
Monod, 1942	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)}$	2.3
Teissier, 1942	$\mu = \mu_{max} \cdot \left(1 - e^{-\frac{S}{KS}}\right)$	2.4
Moser, 1958	$\mu = \frac{\mu_{max} \cdot S^n}{(Ks + S^n)}$	2.5
Contois, 1959	$\mu = \frac{\mu_{max} \cdot S}{(Ks \cdot X + S)}$	2.6
Powell, 1967	$\mu = \frac{\mu_{max} \cdot S}{(Ks + K_D + S)}$	2.6
Mason & Milles, 1976	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)} + K_D \cdot S$	2.7
Vavilin, 1982	$\mu = \frac{\mu_{max} \cdot S^n}{(Ks^{n-P} S^P + S^n)}$	2.8

2.4.1.2. Substrate-Independent Growth Kinetics. In certain cases the application of substrate-independent kinetics given in Table 2.11 can make sense such that the growth rate does not mainly depend on a limiting substrate. For example, in cultivations of mammalian

cells on micro-carriers the cells need free surfaces to colonize and grow. Growth stops if no unoccupied surface is available. Such kinetics can be interpreted as an inhibition by the cell concentration or a space limitation. For this type of fermentation, the kinetics of Frame and Hu can give better description than the logistic law because it exhibits a sharper transition to the stationary phase. Another application of these kinetics is for processes in which information on the limiting substrate is not available, e.g., the limitation of an unidentified component in complex media (Rehm et al., 1991).

TABLE 2.11. Substrate independent growth kinetics (Rehm et al., 1991)

Model Name, Year Proposed	Kinetic Model	Equation Number
Logistic Law, 1938	$\mu = \mu_{max}\left(1-\frac{X}{X_{max}}\right)$	2.9
Frame & Hu, 1988	$\mu = \mu_{max}\cdot\left(1-e^{-\frac{K(X_{max}-X)}{X}}\right)$	2.10

2.4.1.3. Substrate and Product Inhibition Kinetics. Besides substrate limitation, inhibition by substrates or products is quite often found in biotechnological processes. Table 2.12 gives a list of inihibition kinetics. Most of the applied kinetics are extension of Monod equation and have been derived from enzyme inhibition kinetics.

Andrews analyzed substrate inhibition using the equation proposed by Haldane in 1930 to describe enzyme inhibition. by the formation of inactive complexes of the enzyme more than one substrate molecules (Rehm et al., 1991). Edwards proposed a model by modifying also Haldane kinetics (Rehm et al., 1991). The equation has a four parameter. The model implies indefinite cell growth. A quite different equation for the quantification of substrate inhibition kinetics was proposed by Tseng and Wayman. A threshold substrate concentration below which the organism grows apparently without inhibition and above which linearly inhibits the growth. Aiba developed an empirical approach for the product

inhibition patterns of yeast. Edwards suggested to use the same relation to describe substrate inhibition. Edwards also proposed a modified version of the equation that was initially derived by Teissier. In this three-parameter model, the assumption of diffusion-controlled substrate supply has been combined with substrate limitation at high and inhibitory concentrations.

TABLE 2.12. Model functions for inhibiton kinetics

Name	Kinetic Model	Equation Number
Andrews, 1968	$\mu = \frac{\mu_{max} \cdot S}{(K_S + S) \cdot \left(1 + \frac{S}{K_i}\right)}$	2.11
Edwards, 1970	$\mu = \frac{\mu_{max} \cdot S}{K_S + S + \left(\frac{S^2}{K_i}\right) \cdot \left(1 + \frac{S}{K_P}\right)}$	2.12
Tseng and Wayman, 1942	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)}$ when S<S* $\mu = \frac{\mu_{max} \cdot S}{(Ks + S)} - K_i(S - S^*)$ when S<S*	2.13
Aiba, 1968	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)} \cdot e^{-\frac{P}{Ki}}$	2.14
Edwards, 1969	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)} \cdot e^{-\frac{S}{Ki}}$	2.15
Edwards, 1970	$\mu = \mu_{max}\left(e^{-\frac{S}{Ki}} - e^{-\frac{S}{K_S}}\right)$	2.16
Luong, 1987	$\mu = \frac{\mu_{max} \cdot S}{(Ks + S)} \cdot \left(1 - \frac{S}{Sm}\right)^n$	2.17

Luong described another model for correlating the substrate inhibiton (Luong, 1987). The model equation was originally derived by Levenspiel (Rehm et al., 1991) for the

influence of ethanol production on the rate of alcoholic fermentation.The four-parameter model is of the generalized Monod type and accounts for both substrate and stimulations at low S_0 and substrate inhibition at high S_0. The model can be used to predict the maximum substrate concentration above which growth is completely inhibited. The magnitude of n will indicate the type of relation between μ and S. As n approaches unity, model becomes equivalent to the $S/Ki<<1$ case in Equation 2.17. A rapid initial drop in the growth rate followed by a slow decrease occurs when $n>1$ and a slow decrease in the growth rate followed by a rapid decrease occurs when $n<1$.

2.4.2. Structured Models

In structured models, additional variables that accounts for the changes in the micoorganism's composition are used. One of the problem associated with the construction of structured models is that of complexity. In principle an extensive number of compositional variables can be attributed to the biomass. Specifically these models genarally contain a large number of parameters which need to be estimated in critical experiments.

A class of potentially useful models results from a simple extension of the unstructured approach in which the amount and the properties of the biomass are specified by up to three or four variables. Those models are generally termed **compartment models** and combine a better description of the behaviour of the system with moderate mathematical complexity and relatively low number of variables in order to permit experimental verification. In fact, structured models are divided into many groups. **Metabolic models**, are more specific to a particular organism or process since the models include more biological details on the cell metabolism (Bailey and Ollis, 1986). **Genetically structured models** are based upon molecular mechanisms for protein production and combined with the population balance model to predict plasmid-encoded gene product formation, then the model is extended it to the reactor conditions (Laffend and Shuler,

1994). In **chemically structured models**, product formation kinetics have been described chemically.

In **compartment models**, the biomass is compartmentalized into a small number of components. Early attemps to include structure in the description of the biomass were based on the distinction of two compartments in the biomass (Roels,1983). The distinction between these parts was based on a variety of subdivision of the biomass in separate parts. One approach distinguishes between a compartment, responsible for the synthesis of the cellular macromolecules, a synthetic part, and a structural section containing the macromolecules necessary for the functioning of the cellular machinery.

In the compartment model used by Williams et al (Roels,1983), the biomass is divided into two compartments: K and G. The K compartment is assumed to be synthesized from substrate S. The rate of substrate consumption is given by the following kinetic equation:

$$r_{SK} = k_1 \, x \, s \qquad (2.18)$$

The G compartment of the biomass is assumed to be synthesized from K-compartment. The rate of K-compartment consumption is given by the following kinetic equation:

$$r_{KG} = k_2 \, K \, G \qquad (2.19)$$

where K and G are the concentrations of the K and G compartments expressed per unit culture volume.

Roels and Kossen modified the model proposed by Williams (Roels,1983). Their modification involved the stoichiometry of the processes of K and G compartment synthesis, kinetic rate equation for the uptake of substrate and the allowance for maintenance processes. In the models of Williams, direct proportionality between the rate of substrate uptake and the substrate concentration is assumed, i.e. the net rate of substrate

uptake is first order with respect to x. Roels et al (Roels, 1983) have used the following kinetic equation for the rate of uptake of substrate:

$$r_s = q_{max}\ s\ x / (K_s + s) \qquad (2.20)$$

where q_{max} is the maximum specific rate of substrate consumption. Also they have used yield factors to formulate the conversion of the process as follows:

$$S \text{ -------}> Y_{sK}\ K$$

$$K \text{ -------}> Y_{GK}\ G$$

and they have suggested the following alternative for the rate equation:

$$r_K = k_2\ x_K\ x_G\ x \qquad (2.21)$$

$$r_G = m_G\ x_G\ x \qquad (2.22)$$

where m_G is the specific rate of maintenance.

Nielsen et al (1989, 1991) have proposed a four compartment model to describe the microbial kinetics. In their model , biomass is divided into four compartments called A,P , E and G. The A compartment is the active part of the cell and it comprises ribosomes, mRNA, tRNA and various activated building blocks. Since for microorganisms the RNA portion of the ribosomes is almost constant at about 60%, A compartment was estimated by measurement of RNA. P compartment comprises only the plasmid DNA, and E compartment contains only major plasmid product. The rest of the cell was contained in the G compartment which represents the genetic and the structural part of the cell. Only one limiting substrate was used and microbial and reactor kinetics were decoupled by introducing the intracellular component S which is in equilbrium with the extracellular component s (glucose). The compartments P, E, G are built from the copartment A. The reaction kinetics are as follows:

S ------->γ_{11} X_A $$r_1 = \frac{k_1 \cdot S \cdot XA}{S + K_1}$$ (2.23)

X_A ------->γ_{22} X_G $$r_2 = \frac{k_2 \cdot S \cdot XA}{S + K_2}$$ (2.24)

X_A ------->γ_{33} X_P $$r_3 = \frac{k_3 \cdot S \cdot XA}{S + K_3} \cdot \frac{X_P}{X_{P,n}}$$ (2.25)

X_A ------->γ_{44} X_E $$r_4 = \frac{k_4 \cdot S \cdot XA}{S + K_4} \cdot \frac{X_P}{X_{P,n} + K_P}$$ (2.26)

Where $X_{P,n}$ is the normal plasmid content and k_i and K_i are the stauration and stoichiometric constants. X_i denotes the fraction of compartment (g / gcell). They have taken the stoichiometry and the saturation constants of the reactions formed from A compartment as identical ($K_2 = K_3 = K_4$ and $k_2 = k_3 = k_4$).

They have this model to describe runaway plasmid replication. The plasmid replication is controlled at low temperatures and independent of the actual copy number, only a certain number of plasmids is formed in each cycle. There the kinetics in Equation (2.37) was modified to

$$r_3 = \begin{cases} \dfrac{k_3 \cdot S \cdot X_A}{S + K_3} & T < T_{crit} \\ \dfrac{k_3 \cdot S \cdot X_A}{S + K_3} \cdot \dfrac{X_P}{X_{P,n}} & T > T_{crit} \end{cases}$$

The mass balance for the four compartment model are

$$\frac{dX}{dt} = \Gamma^T - \mu \cdot X \qquad (2.27)$$

where the specific growth rate is calculated by

$$\mu = (1)^T \Gamma^T \mathbf{r} = \gamma_{11} r_1 - (1 - \gamma_{22})(r_2 + r_3 + r_4) \qquad (2.28)$$

where the stoichiometric matrix is equal to

$$\Gamma = \begin{bmatrix} \gamma_{11} & -1 & -1 & -1 \\ 0 & \gamma_{22} & 0 & 0 \\ 0 & 0 & \gamma_{22} & 0 \\ 0 & 0 & 0 & \gamma_{22} \end{bmatrix}$$

The simulation results obtained by Nielsen et al agree well with the experimetal results. They have also used compartment model concept to investigate lactic acid fermentation when two-three substrate were used (Nielsen et al., 1989 ;Nielsen et al., 1991a; Nielsen et al., 1991b).

2.4.3. Segregated Models

The continuum approach to the desriction of the behaviour of the cultures of organisms assumes a homogenous distribution of the organisms throughout the culture. However, microorganisms are cellular in nature and the continuum description is not rigorously correct. Sometimes continuum approah can be derived from a segregated approach by averaging techniques (Cazzador, 1991).

The development of a segregated model involves the specification of a set of equations from which the time dependence of the distribution function and the probabilty density function can be calculated. Then constructing the balance equation for the number of elements present in the volume follows by essentially two types of processes. One is the transport of elements to the volume and the other one is the conversion of objects in the volume element (Roels, 1983).

In 1985, Seo and Bailey (1985) have proposed a segeregated, population balance model for obtaining plasmid content distribution based at the single cell level. They considered that the individual cells containing different numbers of plasmids were expected to have different probabilites of plasmid loss at division. They also obtained the population product synthesis activity in unstable recombinant organims that segregate their volume and mass by binary fission upon division.

They have derived the distribution of states W(n) by first determining the age distribution of the cells and subsequently converting age distribution to a plasmid distribution based upon regulation of plasmid replication and segregation at the single cell level.

$$\frac{dW(a)}{da} = -\mu \cdot W(a) \qquad (2.29)$$

where a denotes the cell age since birth and W(a) is the cell age distribution. They have obtained a cell balance relating dividing and new born cells as follows

$$W(0) = (2 - \theta)\, W(P) \qquad (2.30)$$

where p denotes the dividing cell age and θ is the probability of birth of a plasmid free cell upon division of a plasmid containing cell. And using the above equation they have obatined the a relation between specifc growth rate and the probability of birth of a plasmid free cell from plasmid containing cell as follows:

$$\mu = \ln(2 - \theta) / P \qquad (2.31)$$

Then they have proposed another equation for plasmid distribution at birth and using this together with age distribution, they have obtained the distribution of single cell plasmid content in the population after specification of plasmid replication kinetics.

3. MATERIALS AND METHODS

3.1 Materials

3.1.1 Bacterial Strains and Plasmids

Escherichia coli strain 294 containing the plasmid pPG430 which is a derivative of pBR322 and *Escherichia coli strain M5248* (λ+ bio 275 cI857 HI) which was transformed by the plasmid pSCC2 were used for the production of *Eco*RI in this study. Both of the recombinant plasmids carrying the *Eco*RI endonuclease and methylase gene confer resistance to ampicillin.

Escherichia coli strain 294 (pPG430) was kindly provided by Dr. Herbert Boyer (Department of Biochemistry, University of California, San Francisco*). Escherichia coli strain M5248* (λ+ bio 275 cI857 HI) , *Escherichia coli strain N99* (λ+ str+ su-) and plasmid pSCC2 were kindly provided by Paul Modrich (Department of Biochemistry, Duke University- Medical Center, Durham, North Carolina).

3.1.2 Bacterial Culture Media

The Media used throughout this study were LB medium, supplemented LB medium and M9 minimal medium.

LB Medium

Tryptone	10g
Yeast extract	5g
NaCl	10g

per liter of deionized and distilled water.

Supplemented LB Medium

Tryptone	10g
Yeast extract	5g
NaCl	10g
Thiamine	5mg
Thymine	10mg
K-phosphate(pH 7.4)	0.01M
Glucose	5g

per liter of deionized and distilled water.

LB Agar Medium

Agar	15-20 g

per liter of LB Medium

M9 Minimal Medium

M9 Salts (10x)	100ml
1M $MgSO_4.7H_2O$	1ml
1M $CaCl_2$	1ml
L-Leucine	20mg
L-Proline	20mg

per liter of deionized and distilled water.

M9 Salts(10x)

Na_2HPO_4	60g
KH_2PO_4	30g
NaCl	5g
NH_4Cl	10g

per liter of deionized and distilled water.

The M9 minimal medium was supplemented with 1,5,10,15,20 and 30 g/L glucose when necessary.

Selective media were prepared by addition of 80-100 μg/ml of ampicillin to either the liquid medium or the agar medium after they were cooled down to 50°C following sterilization.

3.1.3 Chemicals

All chemicals and solutions used in this study were purchased from MERCK (GERMANY) or SIGMA (USA) unless stated otherwise in the text.

The restriction enzymes were from PROMEGA(USA) and from New England Biolabs(USA). The instant films (667) were from POLAROID (USA).

3.1.3.1 DNA Size Marker Used in Electrophoresis.

λ-DNA/ HindIII DNA Size Marker	23 130, 9416, 6557, 4361, 2322, 2027, 564, 125bp

3.1.3.2 SDS Molecular Weight Markers.

Albumin, Bovine	66kDa

Albumin, Egg	45kDa
Carbonic Anhydrase	29kDa

3.1.4 Buffers and Solutions

3.1.4.1 Assay Buffer.

Buffer A

K-phosphate (pH 7.0)	20mM
2-mercaptoethanol	10mM
EDTA	1mM
TritonX-100	0.2%

20mM K-phosphate was prepared by adding 20mM KH_2PO_4 to 20mM K_2HPO_4 until the pH of the solution was neutral.

3.1.4.2 Lysis Buffers.

Lysis Buffer I

NaCl	0.8M
PMSF	20mM

in Buffer A.

Lysis Buffer II

Tris-HCl (pH 8.0)	10mM
NaCl	0.1M
EDTA (pH8.0)	1mM
TritonX-100	5%
Lysozyme(in 10mMTris.HCl-pH8.0)	0.25mg

3.1.4.3 Dialysis Buffer.

NaCl	0.4M
in Buffer A.	

3.1.4.4. Storage Buffer.

Glycerol	50%
BSA	10mg/ml
NaCl	0.5M
K-PO_4 (pH 7.0)	10mM
2-mercaptoethanol	5mM
TritonX-100	0.1%

3.1.4.5 Agarose Gel Electrophoresis Buffers.

5XTBE (Tris-Borate) Buffer

Tris-Base	445mM
Boric Acid	445mM
EDTA	10mM

10X Loading Buffer

Bromphenol Blue	2.5mg
SDS	1% v/v
Glycerol	1ml

3.1.4.6 Polyacrylamide Gel Electrophoresis Buffers.

Acrylamide-bisacrylamide Mixture (30:0.8)

Acrylamide		29.2g
N'N'-bis-methylene-acrylamide		0.8g
dH_2O	up to	100ml

Sample Buffer

0.5M Tris-HCl pH6.8		1ml
Glycerol		0.8ml
10% (w/v) SDS		1.6ml
2-mercaptoethanol		0.4ml
0.05ml% (w/v) Bromophenol blue		0.2ml

5X Running Buffer

Tris-base		15g
Glycine		72g
SDS		5g
dH_2O	up to	1L

Electrophoresis Separating Gel (10%)

1.5M Tris-HCl pH8.8		2.5ml
Acrylamide/Bis(30:0.8)		3.33ml
10% (w/v) SDS		100µl
10% Ammonium persulphate		50µl
TEMED		5µl
dH_2O	up to	10ml

Electrophoresis Separating Gel (7.5%)

1.5M Tris-HCl pH8.8		2.5ml
Acrylamide/Bis(30:0.8)		2.5ml
10% (w/v) SDS		100µl
10% Ammonium persulphate		50µl
TEMED		5µl
dH_2O	up to	10ml

Electrophoresis StackingGel (5%)

1.5M Tris-HCl pH6.8		2.5ml
Acrylamide/Bis(30:0.8)		1.7ml
10% (w/v) SDS		100µl
10% Ammonium persulphate		50µl
TEMED		10µl
dH_2O	up to	10ml

Stain

Coomasive blue R-250	0.1%
Methanol	40%
Acetic Acid	10%

Destain

Isopropanol	25%
Acetic Acid	10%

3.1.4.7 Buffers and Solutions used in Plasmid DNA Isolation.

STET Buffer

Tris-HCl pH 8.0	10mM
NaCl	100mM
EDTA	1mM
TritonX-100	5%

STE Buffer

Tris-HCl pH 8.0	10mM
NaCl	100mM
EDTA	1mM

Solution I

Tris-HCl pH 8.0	25mM
D-Glucose	50mM
EDTA	10mM

Solution II

NaOH	0.2N
SDS	1%

Solution III

5M Potassium Acetate	60ml
Glacial Acetic Acid	11.5ml
dH_2O	28.5ml

TE Buffer

Tris-HCl pH 8.0	10mM
EDTA	1mM

Phenol TE saturated, pH 8.0

Chloroform / Isoamylalcohol 24:1(v / v)

Ammonium acetate 3M

Cracking 2XBuffer

5M NaOH	2ml
10% SDS	2.5ml
Sucrose	10g

dH_2O	up to	50ml

3.1.4.8 Solutions used in RNA Estimation.

Ferric Chloride Reagent

$FeCl_3.6H_2O$	100mg
Dissolve in concentrated HCl	100ml

Ethanolic Orcinol Reagent

Orcinol	6.0g
Ethanol	100ml

3.1.5 Laboratory Equipment

Autoclaves	Medexport, CIS, C.W.I.S. Eyela, Model MAC-601, Japan
Balances	Precisa 80A-200M, Switzerland
Camera	Polaroid DS-34 Direct Screen Instant Camera, USA
Centrifuges	SORVALL RC-5B Refrigerated Superspeed Centrifuge, DUPONT, USA Biofuge 28RS HERAUS, Germany Centrifuge 5415 C EPPENDORF, Germany Centrifuge NF 615, NÜVE, Turkey

Cold Room	VWR Scientific, VCR 422DBA, USA
Deepfrezers	-80°C, Hetofrig CL 89,Heto, Denmark -20°C, BOSCH, Germany -20°C, ARÇELİK, Turkey -20°C, Ultra 1500, SIMTEL, Turkey
Electrophoresis	Horizon58, Model200, Horizontal Gel Electrophoresis Apparatus, BRL, USA Miniprotean II, BIORAD, USA Power / Pac 300, BIORAD, USA
Fraction Collector	Model 2110, BIORAD, USA
Freeze Drier	Chem Lab Instruments Ltd, Model SB6, England
Gel Dryer	Heto Dry GD-I, Denmark
Hot Plate	Nüve 318, Turkey
Laminar Hood Cabinet	Holten-Lamin Air, HBB 2460, Denmark
Ice Machine	Scotsman, AF-30, UK
Incubators	EN500, NÜVE, Turkey FN500, NÜVE, Turkey Lab-Line AMBI-HI-Low Chamber, 3554-18, USA
Magnetic Stirrer-Heater	MK 318, Nüve, Turkey
Membrane Pump	Vacuubrand, ME2, Germany

Orbital Shaker	GFL 3032, Germany Innova, 4340, New Brunswick Scientific Co., England
pHmeter	HANNA Instruments, HI 8521, Singapore
Recorder	Model 1325, Econo Recorder, Biorad, USA
Refrigator	+4°C, ARÇELİK, Turkey +4°C, SİMTEL, Turkey
Sonifier	Model 250/450 Sonifier Branson Ultrasonic Corporation, USA
Spectrophotometer	DU 640 Beckman, USA
Thermo-cycler	Thermal Reactor TR1, HYBAID, UK
Transilluminator	Reprostar II, CAMAG, Switzerland
UV Monitor	Econo UV Monitor, Biorad, USA
Vortex	Elektro-mag, Turkey
Water Baths	Heto, CB 8-30e, Denmark Heto, DT Hetotherm, Denmark Nüve, BM 102, Turkey
Water Distillation System	Millipore, Milli Ro Plus,USA Millipore, QVF Plus, USA GFL-2004, GERMANY

3.2 Experimental Methods

Sterilized equipment was used during all experiments. Pipetman tips, eppendorf tubes, centrifuge tubes, columns, solutions and culture media were all sterilized at 1.02 atm and 121°C for 20 minutes in an autoclave. All glassware were sterilized at 180°C for 2 hours in an oven.

The strains of *E.coli* were kept in glycerol solution at -70 °C. Frozen glycerol cultures were used by streaking on LB agar plates for the preparation of master plates which were used to remove one isolated colony when it was required.

3.2.1 Preparation of Pre-culture

50 ml of sterile liquid nutrient medium was inoculated with a single colony of bacteria from a slant culture. The inoculation was performed by an inoculating needle which was flamed to sterilize and cooled on the agar plate. The preculture was incubated in an orbital shaker at 30-32°C in the case of *E.coli N99/λ+ and E.coli M5248* cells or at 37°C in the case of *E.coli 294* cells to an OD_{600} of approximately 0.6. The liquid and the agar mediums were supplemented with 80-100 μg/ml ampicillin concentration.

3.2.2 Large-Scale Plasmid Isolation by Alkaline Lysis Method

Large-scale plasmid isolation by the alkanline lysis method was carried out as described in Maniatis (1984). 500 ml of LB medium containing ampicillin was inoculated

with 15 ml of the pre-culture of *E.coli N99/λ+* cells carrying plasmid pSCC2 and incubated for 16-24 hours at 32°C with vigorous shaking.

Bacterial cells were harvested by centrifugation at 4000 rpm for 15 minutes at 4°C, resuspended in 100ml STE buffer and collected by centrifugation. 18 ml Solution I, 100 mg lysozyme and 40 ml freshly prepared Solution II were added to the bacterial pellet for 20 minutes of incubation at room temperature. Then, 20 ml ice-cold Solution III was added to the mixture, and it was incubated on ice for 15 minutes by gentle mixing by inversion. The lysate was then centrifuged at 6000 rpm for 15 minutes at 4°C and the supernatant was filtered into a sterile centrifuge tube.

0.6 volume of isopropanol was added to the supernatant to concentrate the plasmid DNA contained in the cleared lysate. DNA solution was then mixed by inversion and incubated 30 minutes at room temperature. Nucleic acids were removed by centrifugation at 1000 rpm for 15 minutes at room temperature, the pellet was washed with 70 % ethanol and dissolved in 10 ml TE Buffer by gentle pipetting.

DNAase-free pancreatic RNAase was added to this solution to a final 10 mg/ml concentration and incubated one hour at room temperature. One volume of phenol:chloroform (1:1) was added to the solution and shaken vigorously to remove the proteins. After centrifugation at 6000 rpm for five minutes, the upper phase was mixed with one volume of chloroform:isoamylalcohol (24:1). Plasmid DNA was precipitated by the addition of 0.3 M ammonium acetate and 2.5 volumes of absolute ethanol.

The final mixture was kept at -70°C for one hour or at -20°C overnight, the plasmid DNA was recovered by centrifugation at 15000 rpm for 30 minutes at 4°C. The pellet was washed with 70% ethanol and dried. 500 µl sterile TE Buffer was used for the final resuspension and DNA was stored at -20 °C for long term use.

3.2.3 Transformation of *E.coli* M5248 Cells by the Plasmid pSCC2 by $CaCl_2$ Method

Transformation of *E.coli* M5248 cells by the plasmid pSCC2 by $CaCl_2$ method was carried out as described in Maniatis (1984). An overnight pre-culture of E.coli M5248 cells in LB medium was prepared at 37°C. 100ml LB medium was inoculated with one ml of the overnight culture and incubated at 37 °C to an OD_{550} of 0.5. Then the cells were cooled for 15 minutes on ice and pelleted by centrifugation at 4000 rpm for 10 minutes at 4°C by SS-34 type rotor.

The cell pellet was suspended in 40 ml of 0.1 M $CaCl_2$ and stored on ice for 30 minutes. The suspension was centrifuged again at 4000 rpm for 10 minutes at 4°C by SS-34 type rotor, and the pellet obtained resuspended in 4 ml of 0.1 M $CaCl_2$ and incubated for two hours on ice.

The competent cells were separated into 200 µl aliquots and were placed in sterile microtubes by chilled and sterile pipette tips. 10-40 µg of the plasmid pSCC2 was added to the tubes and the microtubes were stored on ice for 30 minutes after mixing by gentle swirling. One microtube containing no DNA was used as negative control of the transformation. The cell suspension was heated for three minutes in a 32°C water bath. 800 µl of preheated LB medium was added to the suspension. The cells were incubated at 32°C for two hours to express the ampicillin resistance marker encoded by pSCC2.

The different volumes of the transformants together with negative control were spread on LB agar plates containing 100 µg/ml ampicillin concentration for selection of ampicillin resistant cells. The plates were left at room temperature to diffuse for about half an hour and then incubated at 32°C for approximately 20 hours at an inverted position. No colonies were seen on negative control plates.

3.2.4 Cracking Procedure for Rapid Plasmid Detection in *E.coli M5248*

The colonies of transformed *E.coli M5248* cells were checked to see whether they still carry the plasmid pSCC2 before being pre-cultured. The presence of plasmid in transformed cells was shown by using a procedure described by Rodriguez and Tait (1983).

The colonies were dissolved in 25 µl 10 mM EDTA (pH 8.0) and 25 µl of freshly prepared cracking 2X buffer was added. The mixtures were incubated at 70°C for 10 minutes and left to cool down to room temperature. 1.5 µl of 4M KCl and 0.5µl of 0.4% bromophenolblue were added to the samples which were vortexed and incubated for 10 minutes on ice. Then they were centrifuged at 14000g for three minutes and 10-20 µl of the supernatant was immediately loaded on an agarose gel. Plasmids of known size were loaded on the same gel as markers.

3.2.5 Determination of the Population Growth Curve

Ten ml of pre-culture was used to inoculate sterile 1L LB medium and then placed in the orbital shaker at the specified temperatures for each *E.coli* strain. An aliquot of the cell suspension was taken at desired time intervals. The optical density values of samples were measured at 600nm using cell-free broth as the blank. For the estimation of optimum growth period, a calibration curve of absorbance versus dry cell mass was constructed.

3.2.6 Harvesting Cells

Following the growth of *E.coli 294* strain to an absorbance value of 1.2 at 590nm at 37°C, the cells were induced by the addition of 0.1mM IPTG concentration with an induction period of 6 hours.

When the cell concentration of *E.coli M5248* strain reached an absorbance value of 1.0 at 590nm at 30°C, the culture temperature was raised to 42°C and the incubation was continued over four hours to induce the cells.

After the growth period was completed for each *E.coli* strain, the cultures were divided into centrifuge tubes. The cells were then harvested by centrifugation at 4000rpm for 30 minutes using GSA type rotor. When the centrifugation of the culture was completed, cells were resuspended in 0.9% ice-cold NaCl solution and pelleted again at the same conditions. At this stage, cells were ready for purification.

3.2.7 Purification of *Eco*RI Endonuclease

The purification of the restriction endonuclease *Eco*RI was accomplished by using the following procedure that was developed by modifying and improving the method available in the literature (Greene, 1974; Luke and Halford, 1985). All steps of the purification were performed at 0-4°C.

3.2.7.1. Preparation of the Cell Lysate. For transferring the intracellular enzyme into the surrounding solution, the rupture of the cell membrane is necessary. In this study, two different methods were used for the disruption of the cell membrane: (i) the enzyme lysozyme was used to digest the cell wall for the 1.5ml aliquots taken at regular time

intervals, (ii) the sonic oscillation method was used to disrupt the cell membrane of the culture with larger volumes. When suspensions are subjected to ultrasonic vibrations, the high transient pressure produced breaks the cell wall.

(i) by addition of lysozyme; the bacterial pellet was obtained from 1.5ml culture by centrifugation at 10000g for 2 minutes. The pellet was suspended in 350µl of STET buffer, and 25µl of freshly prepared 10 mg/ml solution of lysozyme prepared in 10mM Tris.HCl (pH 8.0) was added to the solution. The samples were incubated in boiling water bath for exactly 40 seconds. The bacterial lysate was centrifuged at 12000g for 10 minutes at room temperature. The supernatant was removed to a sterile microcentrifuge tube to measure the enzyme activity.

(ii) by ultrasonication; the cell pellets were suspended in sonication buffer with a ratio of 2.5ml/100ml of cell culture. In order to minimize the effect of the heat generated, the flask containing the suspended cells was cooled by placing it in ice. Furthermore, during the sonication the cell suspension was circulated away from the ultrasonic probe to minimize the localized heating effects. The cells were sonicated for 5 x 30 seconds.

After the sonication was completed, the cell extract was replaced in the dialysis tubes. The dialysis tubes were prepared by submerging the required length of tubing in a solution of 5% (w/v) sodium carbonate and 1mM EDTA and then sterilizing in the autoclave for 30 minutes. The tubing was washed out with distilled water before the extract was transferred into it. The ends of the bag was squeezed to prevent any leakage and the tubing was placed in dialysis buffer which was mixed continuously for 16 hours in the cold room.

At the end of the dialysis period, cell debris and subcellular materials were removed by centrifugation at 9500 rpm for 15 minutes by using SS-34 type rotor. The clarified lysate was saved for purification of the enzyme by successive applications of phosphocellulose and hydroxyapatite chromatography.

3.2.7.2 Application of Phosphocellulose Chromatography. 62.5 grams of phosphocellulose was suspended in two 2 liters of 0.1N HCl-48% ethanol and stirred for 30 minutes at room

temperature. The slurry was collected by vacuum filtration on a Buchner-funnel and then resuspended and stirred into two liters of distilled water. This step was repeated until the pH of the slurry was near neutral. The phosphocellulose was then collected by filtration and resuspended in two liters 0.1N NaOH and mixed at room temperature for 30 minutes. The phosphocellulose was filtered again and resuspended in two liters of 1mM EDTA and stirred for 30 minutes. The slurry was refiltered and washed with distilled water until its neutral pH.

The phosphocellulose was finally equilibrated in Buffer A and the pH was adjusted to pH 7.0 by using 5N HCl. The packing of the 2x30 cm column was carried out by gently pouring the slurry of the phosphocellulose into the column. During the packing, it was ensured that the slurry was poured from the edge of the column so that no air bubbles were trapped and any disturbance that can occur on the surface was prevented. The slurry was added until the required height calculated according to the five ml phosphocellulose/gram cell ratio was obtained and the column was transferred to the cold room. After the slurry in the column settled down, the column was washed with minimum two times the packed volume of Buffer A + 0.4M NaCl.

The cell lysate was applied to the equilibrated column carefully by allowing it just to run slowly into the column from the edges. Following the application of the sample, the effluent was collected in a sterile flask for further testing. The subsequent elution was carried out stepwise by buffer A containing increasing concentrations of NaCl (0.4, 0.5 ,0.6, 0.7, 0.8, 1.0M). During each elution, one-three ml fractions were collected by using a fraction collector. The time for shifting to a higher salt concentration was decided according to the elution profile which was obtained by reading the optical density at 280nm. Each fraction within the peaks of the elution curve was assayed for *Eco*RI endonuclease activity (Section 3.3.5). Active fractions, which eluted at about 0.6M NaCl, were pooled and applied to hydroxyapatite chromatography for further purification.

3.2.7.3 Application of Hydroxyapatite Chromatography. One ml of hydroxyapatite solution (BIORAD) was used per gram cell. The slurry of hydroxyapatite was washed several times with Buffer A, then equilibrated in Buffer A + 0.6M NaCl.

The pooled endonuclease fractions were applied to batchwise hydroxyapatite chromatograpy. After a 30-minutes mixing period in the cold room, the slurry was centrifuged at 5000rpm for 15 minutes by using SS-34 type rotor. The supernatant was saved to be tested. The pellets were washed with increasing concentrations of K-phosphate ranging from 0.1M to 0.6M in Buffer A with 0.6M NaCl. After suspending the pellets in each buffer solution, the slurry was centrifuged at the same conditions following a 15 minutes mixing period in the cold room. The fractions containing active enzyme were pooled.

3.2.7.4 Application of Second Phosphocellulose Chromatography. The pooled enzyme solution obtained from hydroxyapatite chromatography was diluted four times with Buffer A and immediately applied to a second phosphocellulose column (1x10 cm). The elution was carried out as in the first phosphocellulose column (Section 3.2.7.2). The active fractions determined by the enzyme assay were pooled

3.2.7.5 Storage. The final enzyme solution was supplemented with 50 μg/ml BSA and dialyzed against storage buffer in the cold room for a minimum of 20 hours. After dialysis the resulting enzyme in the glycerol solution was transferred to -20°C.

3.2.8 Application of Quality and Control Tests

3.2.8.1 Cut-Ligate-Recut. *Eco*RI restriction endonuclease was also tested for the presence of contaminants that would inhibit ligation or degrade termini. The restored sites were cleaved by the same enzyme.

After performing the initial cleavage with *Eco*RI isolated in this study, DNA was extracted by phenol and chloroform and precipitated with ethanol. T4 ligase was used to ligate the fragments obtained from the initial cleavage. Ligation was performed at 16°C for 4 hours under conditions described by the manufacturer (Gibco BRL). T4 ligase was

inactivated by heating for 15 minutes at 65°C and the ligation mixture was cooled on ice. Ligated fragments were recut by using the same enzyme preparation. Ligated and the recleaved fragments were verified on agarose gel.

3.2.8.2 Overdigestion Test. Each preparation of *Eco*RI was tested for contamination by the other endodeoxyribonucleases capable of digesting DNA either at random sites or at specific sites appropriate to *Eco*RI. One microgram of substrate DNA is digested with 20 units of enzyme for 5 hours at appropriate temperature. This represents a 100 fold excess digestion as compared to 1 unit for 1 hour as described in the literature (Luke and Halford, 1985; Mehra et al, 1993).

3.2.8.3 Determination of Enzyme Purity. The final enzyme solution was subjected to SDS-PAGE under denaturing conditions by using a slightly modified procedure described by Laemli. (1970).

Denatured samples were loaded into the wells of polyacrylamide gel consisting of five percent stacking and ten percent separating gel. At the end of the electrophoresis, the gel was submerged into the staining solution for 30 minutes. The gel was then submerged into a destaining solution for 45 minutes. Finally the gel was dried and photographed.

3.3 Analysis of Fermentation

3.3.1 Determination of Dry Cell Weight

Dry cell mass measurements were performed by using the vacuum filtration method. Dried and preweighed 0.22μm Millipore filter papers were used to filter the known volumes of samples taken from the culture at regular time intervals. Filtered samples were left to dry

in an oven at 105°C overnight and then weighed. The dry weight of the samples were calculated after subtracting the blank weight of the filter papers.

3.3.2 Determination of Substrate Concentration

The amount of residual glucose concentration in the medium was determined by using D-Glucose Kit of Boehringer Mannheim as described by the manufacturer.

D-Glucose is phosphorylated to glucose -6-phosphate in the presence of the enzyme hexokinase and adenosine-5'-triphosphate(ATP) with the simultaneous formation of adenosine-5'-diphosphate (ADP). In the presence of the enzyme glucose-6-phosphate dehydrogenase, glucose-6-phosphate is oxidized by nicotin amide adenine dinucleotide phosphate(NADP) to gluconate-6-phosphate with the formation of reduced nicotinamide-adenine dinucleotide phosphate(NADPH). The amount of NADPH formed at the end of the reaction is stoichiometric with the amount of D-glucose. The increase in NADPH is measured by means of its absorbance at 340nm.

The test combination contains Solution I; consisting of triethanolamine buffer, NADP, ATP, magnesium sulfate and stabilizers, Solution II; consisting of hexokinase and glucose-6-phosphate dehydrogenase and Solution III as D-glucose standard solution.

A series of cuvettes containing one ml of Solution I which was brought to 20-25°C before use and 0.1ml of diluted sample taken at regular intervals completed with 1.9ml of distilled water were prepared for samples. One cuvette for blank was prepared by addition of 1ml Solution I and 2ml distillate water. The solutions were swirled gently by closing the cuvettes with parafilm and absorbances were read against air as blank, after approximately 3 minutes (A_1).

The reactions were started by addition of 0.02ml of Solution II to cuvettes containing blank and samples. The solutions were again gently swirled and absorbances were read after 15 minutes (A_2). If the reactions have not stopped within the 15 minutes period, absorbances were continued to read at 2 minutes time intervals until the absorbances increase constantly over 2 minutes.

The absorbance differences (A_2-A_1) for both blank and sample were determined and the absorbance difference of the blank was subtracted from the absorbance differences of the samples, thereby obtaining $\Delta A_{D\text{-glucose}}$. The concentration of D-glucose was calculated by the following equation

$$c_{D\text{-glucose}} = 5.441 \times \Delta A_{D\text{-glucose}} \times F / \varepsilon \qquad \text{(g D-glucose/lt sample)}$$

ε: extinction coefficient

F:dilution factor of the samples

3.3.3 Determination of the Protein Content

Bradford Dye-binding Method was used to estimate the protein content of the samples (Bradford, 1978).

A series of tubes containing different dilutions of the samples were completed to a final volume of 800µl with distilled water. Increasing amounts of Bovine Serum Albumin from 0.1 to 60 µg were prepared as standard protein solutions in 800µl of distilled water. 650µl of 1:4 diluted standard Bradford dye reagent (BIORAD) were added to all the tubes, the tubes were mixed several times by gentle inversion. After a period ranging from 5 minutes to 1 hour at room temperature, the absorbance of each sample was read at 595nm against a blank containing 800µl distilled water and 650µl diluted Bradford reagent.

A standard curve showing bovine serum albumin amount in μg protein versus absorbance at 595nm was obtained via a linear regression program. The absorbance values of the samples were converted to μg protein using the calibration curve, the protein concentrations were calculated after dividing by the diluted sample volumes.

3.3.4 Determination of the *Eco*RI Endonuclease Activity

One unit of *Eco*RI endonuclease was defined as the amount of the enzyme that can digest one microgram of λDNA in one hour in a total reaction volume of 10μl.

λDNA was used as the standard DNA of known size and it has five recognition sites for EcoRI enzyme as 21 226, 7421, 5804, 5643, 4878, 3530 bp. Enzyme assays were carried out for each sample obtained from purification. The reactions were performed with one μg λDNA, one μl of 10x reaction buffer and one μl of diluted enzyme solution completed to 10μl of total reaction volume with distilled and sterilized water. The reaction mixtures were incubated at 37°C for one hour. One μl of 100mM EDTA was added to each sample to stop the reaction at the end of one hour and then incubated at 65°C for 5 minutes.

The extent of digestion was analyzed by electrophoresis through 0.8% agarose gels, after the addition of one μl of 10x loading buffer to each sample. At the end of agarose gel electrophoresis, DNA fragments of different lengths were visualized as discrete bands under the UV light and photographed.

3.3.5 Determination of the Intracellular RNA Content by Orcinol Method

The colorimetric assay involving the reaction of aldopentoses with acidified orcinol to produce a green chromogen was used to measure the RNA content. This is a classical reaction of sugar chemistry. The purine-ribose links of RNA are easily hydrolyzed by hot acid and the purine-bound ribose of RNA can be determined colorimetrically. Deoxyribose gives 20% of the color of ribose.

The estimation and identification of ribonucleic acids in cell extracts involves extraction and hydrolysis with a hot solution of perchloric acid followed by specific colorimetric assays for pentose components of nucleic acids. The intracellular RNA was extracted by HCl_4 and the extracted sample was analyzed for ribose by Orcinol Method (Nielsen , 1991c).

Each sample with a volume of 2.5ml of bacterial culture taken at regular time intervals was centrifuged at 10000g for two minutes. The bacterial pellets were acidified by addition of five ml of ice-cold 0.25N HCl_4 and allowed to stand on ice for 30 minutes. The samples were then centrifuged at 10000g for 10 minutes and the precipitates were suspended in four ml of 0.5N HCl_4. After 15 minutes incubation at 70°C with occasional shaking, suspension was centrifuged at room temperature for 10 minutes at 5000g. The cell hydrolysates were diluted with 0.1N HCl so that they contain 10-100 μg of RNA per ml. The diluted perchlorate hydrolysates of the cells were mixed with two volumes of orcinol reagent which was freshly prepared by mixing the equal volumes of ferric chloride reagent with ethanolic orcinol solution. After 30 minutes incubation time at 90°C, the samples were cooled and the absorbance of each sample was determined at 665nm.

The standard curve was obtained by using of Yeast RNA(Sigma). 200 μg of RNA per ml of 0.25N $HClO_4$ was incubated at 70°C for 30 minutes. Different concentrations of RNA were obtained by diluting the sample with 0.1N HCl. The standards were then mixed with two volumes of orcinol reagent and incubated at 90°C for 30 minutes. The calibration

curve was prepared by using linear regression program after reading the absorbances of the standard solutions at 665nm.

3.3.6 Determination of Plasmid Content

The plasmid content was determined by extracting Plasmid DNA from a definite volume of bacterial culture by estimating the plasmid amount spectrophotometrically. The concentration of DNA was calculated as a function of the optical density at 260nm by applying the formula:

$$50\ \mu g/ml \times OD_{260} \times \text{dilution factor} = \text{concentration in } \mu g/ml$$

by taking 50 μg of double stranded DNA has an absorbance of 1.0 at 260 nm.

1.5ml of bacterial culture taken at regular intervals were placed into a micro-centrifuge tube and centrifuged at 10000g for two minutes. The medium was removed and the cell pellet was resuspended in 350 μl of ice-cold STET Buffer.

Microcentrifuge tubes were mixed by vortexing for three seconds after addition of 25μl of freshly prepared solution of lysozyme (10mg/ml of 10mM Tris.HCl- pH 8.0).

Samples were then placed in a boiling water bath for exactly 40 seconds. The bacterial lysate were centrifuged at 12000g for 10 minutes at room temperature, the supernatant was transferred to a sterile microcentrifuge tubes.

40 μl of 2.5M Sodium Acetate (pH 5.2) and 420 μl of isopropanol were added to the supernatant and mixed by vortexing. After incubating the tubes for 5 minutes at room temperature, the nucleic acids were recovered by centrifugation at 12000g for 5 minutes at 4°C.

The pellet was washed with 70% ethanol and recentrifuged at 12000g for 2 minutes at 4°C. All the supernatant was removed by gently aspiration and the pellet was suspended in 50µl of sterile TE (pH 8.0) containing DNAase free RNAase (20µg/ml).

The optical density of the plasmid DNA solution was determined at 260nm against a blank containing only TE buffer. 50µg/ml plasmid DNA was taken as the unit absorbance value at 260nm. The absorbance values were also determined at 280nm to follow the contamination of RNA and protein level in the isolated DNA solution.

3.3.7 Determination of Total Viable Cells

An aliquot of the cell suspension at the desired time intervals was taken to determine the number of viable cells by spread plate technique (Wilson, 1986).

The aliquot taken from culture was serially diluted by the ratio 1:10 in sterile LB medium. The samples were mixed by vortexing before each dilution to achieve homogenous dispersion. 50µl of appropriate dilutions were spread on duplicate LB agar plates (with and without ampicillin) and incubated overnight at the specified temperature for each strain.

Colonies were counted and the number of viable cells per ml of bacterial culture was calculated from the average number of colonies counted on duplicate plates, after correction by the dilution factor.

3.3.8 Determination of the Percentage of Plasmid-Containing Cells

Two different techniques as plate to plate transfer by using a velvet pad and transfer of the colonies by toothpicks, were used to determine the percentage of plasmid-containing cells (Rehm et al., 1991).

The technique of replica plating using a velvet pad has been used for the well isolated colonies on the agar plates. A sterile velvet pad was pressed over a amp^- agar plate containing well separated colonies, the colonies picked by the velvet were transferred on an amp^+ agar plate.

Individual colonies on amp^- LB agar plate were transferred to onto an amp^+ LB agar plate using sterile toothpicks if replica plating was not applicable. Colonies were gently touched with the tips of the sterile toothpicks.

After the amp^+ plates were incubated overnight, colonies were counted. The number of the colonies which can grow on selective medium was divided by the number of the colonies in the first plate with non-selective medium. This ratio obtained was taken as the percentage of the plasmid-bearing cells in the bacterial culture.

4. RESULTS AND DISCUSSION

Restriction enzymes are very important tools of the recombinant DNA technology. Especially, type II restriction endonucleases are very important, since they can cleave DNA at sequence specific sites. A considerable number of highly efficient expression vectors have been engineered in the past few years to obtain high yields of those enzymes encoded by cloned genes of procaryotic origins (Botterman and Zabeau, 1985). Strong controllable promoters are preferred in the construction of the expression vectors to enhance the synthesis of the product. The induction conditions are considered to be important parameters in the production of the recombinant proteins and they need to be carefully optimized to increase the yield of the product.

*Eco*RI restriction endonuclease has proven extremely useful as a reagent in the analysis and manipulation of DNA molecules and therefore it is one of the most widely used type II restriction enzymes. In addition to the natural overproducer of *Eco*RI, *E.coli* RY13, genetically modified overproducing strains were also used to produce enzyme (Greene et al., 1974).

In this study, a parametric study of the fermentation by two different genetically modified *Escherichia coli* strains was conducted to optimize the recovery of *Eco*RI restriction endonuclease.

The first strain used was *E.coli* 294 carrying plasmid pPG430 in which *Eco*RI endonuclease and methylase structural genes were located 220 bp downstream from the lacUV5 promoter. *E.coli* 294 is an overproducer of the *Eco*RI endonuclease. The induction of the *Eco*RI endonuclease was achieved by the addition of IPTG to the medium which was optimized in terms of induction time, induction period and the concentration of IPTG.

The second strain used was *E.coli M5248* carrying plasmid pSCC2, also an overproducer of the *Eco*RI protein with sufficient genetic stability. In this plasmid,

structural genes for *Eco*RI endonuclease and methylase are located about 650 basepairs downstream from the bacteriophage λp_L promoter which is the under control of the cI repressor. *E.coli* M5248 strain fails to produce phage structural proteins upon thermal inactivation of the cI repressor (Cheng et al., 1984). *Eco*RI synthesis is therefore induced by transferring the culture to elevated temperatures. The thermal induction may have several negative effects on the cellular metabolism such as reducing cell growth or even killing the cells. Therefore, it is important to expose the culture to the higher induction temperatures for limited periods of time. In the present work, the induction time and induction period were optimized for the synthesis of *Eco*RI endonuclease by *E.coli M5248* (pSCC2).

In the framework of this thesis, the purification of *Eco*RI endonuclease was accomplished under similar conditions for both strains by using a procedure that was developed by modifying and improving the methods available in the literature (Botterman et al., 1985 ; Cheng et al., 1984; Luke and Halford, 1985). The major steps involved in the purification procedure were cell growth and the preparation of cell extract followed by successive phosphocellulose, hydroxyapatite and phosphocellulose chromatographies. Production yields of *Eco*RI endonuclease from both of these overproducing strains were compared.

The effect of the various culture conditions on the stability of the recombinant plasmids expressing *Eco*RI endonuclease were also studied in both strains by investigating the changes in the growth parameters such as optical density, number of viable cells, dry cell weights and the product yield.

After a study of the effect of substrate concentration on the specific growth rates and *Eco*RI activities of the two recombinant strains, the *E.coli* 294 (pPG430) strain was found to be the more efficient one in producing *Eco*RI endonuclease. Therefore, *E.coli* 294 (pPG430) strain was chosen for investigating the effect of limiting substrate concentrations on cell growth and on each of the intracellular cell components. Experimental data were obtained on the total protein and total RNA concentrations, plasmid content, *Eco*RI endonuclease activity, substrate consumption, total and viable cell mass. These data were used to develop and verify the mathematical models for recombinant protein production under limiting substrate concentrations.

The effect of fermentation conditions on the copy number of the recombinant plasmids were predicted by using the mathematical model evaluated from the simple growth data.

Two different mathematical approaches were used to analyze growth kinetics: unstructured and structured modelling of the biophase. In the unstructured models tested, the microorganism was regarded as a single reacting species with fixed composition and data on the substrate consumption and biomass formation were used. The product concentration was excluded from the models since the enzyme produced is not excreted. In order to construct a structured model for *E.coli* 294 (pPG430), intracellular components such as the total protein, RNA and plasmid DNA contents and *Eco*RI endonuclease activity were measured. A three-compartment model was developed including the RNA concentration, total protein content and *Eco*RI endonuclease activity along with substrate concentration.

4.1 Optimization of Induction Parameters for the Production of *Eco*RI Endonuclease by *E.coli 294* (pPG430)

In *E.coli 294* strain, the structural genes of *Eco*RI are placed in plasmid pPG430 under the control of lacUV5 promoter (Figure 4.1). The induction of *Eco*RI endonuclease was accomplished by the addition of the non-hydrolizable analog of lactose, isopropyl-β-D-thiogalactoside, IPTG.

In order to increase the yield of the *Eco*RI endonuclease from this strain, the induction parameters such as the inducer concentration, the period of induction and the cell concentration when IPTG should be added to the fermentation broth were optimized.

The growth behaviour of recombinant *E.coli* 294 cells containing the overproducing plasmid pPG430 was investigated at 37°C. A lag phase of around three hours was observed

in the growth of cells. The exponential phase of cell growth was completed in 16 hours after inoculation, after which the cells entered their stationary phase. The growth curve is presented in Figure 4.2 together with the four different times of induction used in subsequent experiments as indicated by arrows. The first set of experiments were performed to determine the optimum period or duration of induction; then, the optimum time of induction, i.e. the growth stage at which inducer should be added, was determined in the second set of experiments. 0.1mM IPTG concentration was used throughout the experiments conducted to determine these two parameters.

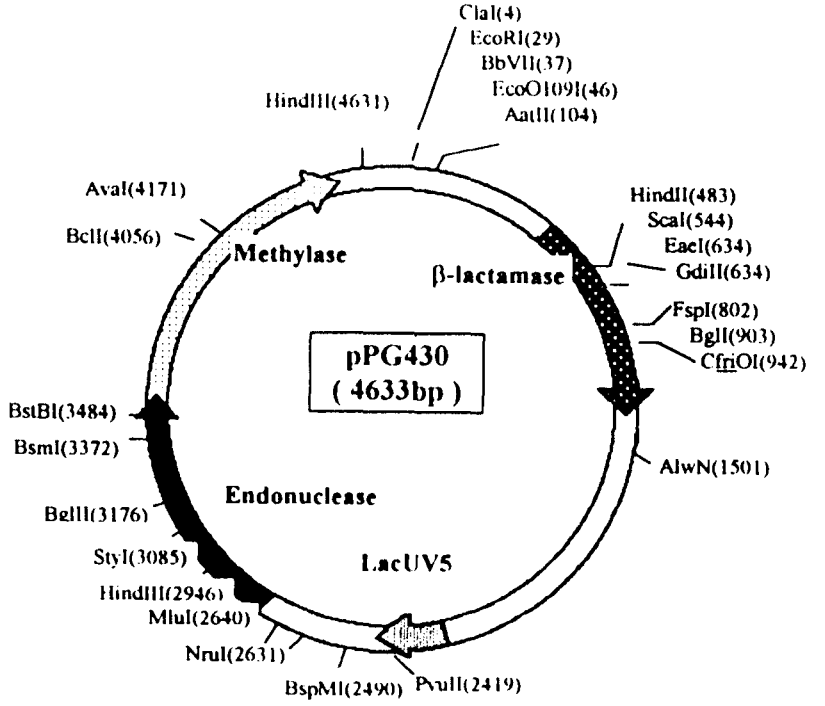


FIGURE 4.1. Structure of plasmid pPG430

In order to determine the *optimum induction period*, two shake flask fermentations were performed in parallel. The first culture was grown at 37°C in LB medium until the absorbance value reached 1.2 at OD_{595}, and then 0.1mM IPTG was added to the medium. The same amount of IPTG was added to the second culture when its absorbance at 595 nm reached 0.5. After the addition of IPTG to the fermentation broths, samples were taken every hour from both culture media to determine the specific activity of the Eco*R*I endonuclease.Cells were distrupted by sonication and the enzyme activities were determined in cell extracts without further purification, to minimize other effects that may arise during

purification. The maximum specific activity of the enzyme was found to be attained after six hours from the addition of IPTG to the medium, then a decrease in specific activity was observed in both of the shake flask fermentations (Figure 4.3). The addition of IPTG in the early or the late exponential phase did not have any meaningful effect on the induction period. Six hours of induction was found to be the optimum induction period for the synthesis of the enzyme.

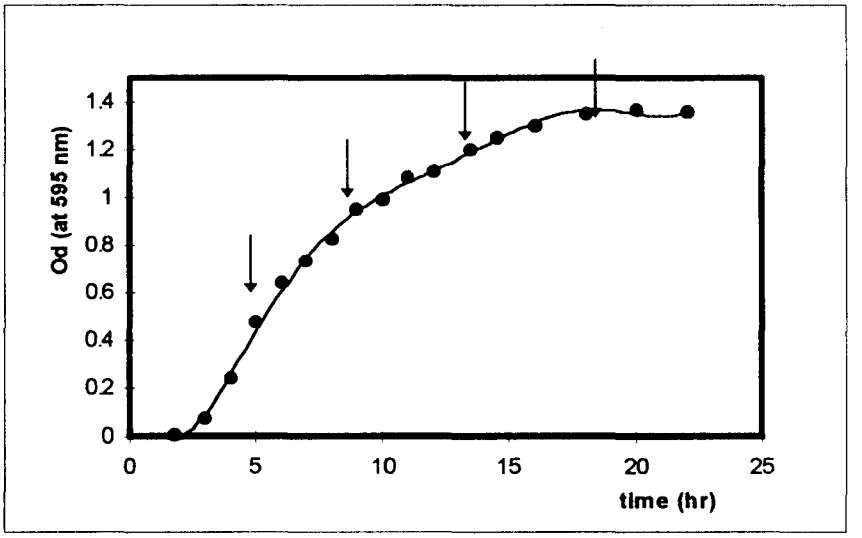


FIGURE 4.2. Growth curve *of E.coli 294* (pPG430) (arrows indicate the induction times at which IPTG was added).

In order to determine the *optimum time of induction*, four shake-flask fermentations were performed in parallel, and each culture was grown to a different absorbance value such as 0.45, 0.9 ,1.2 and 1.4 at 595 nm respectively. Following the addition of 0.1mM IPTG to each of the culture media, the incubation was continued over a 6-hour period, and enzymatic activity was determined in cell extracts. The results are presented in Table 4.1. The specific activity (enzyme produced per mg protein, U/mg protein) and the yield of the enzyme, i.e. the enzyme produced per gram cell (U/gcells), increased as the cell concentration increased until the stationary phase. A decrease in the *Eco*RI enzyme activity was observed if the induction was made at the stationary phase. This observation is in agreement with the results reported by Botterman et al (1985) who showed that an induction of gene expression by a temperature-shift during the exponential phase in the

production of *Eco*RI enzyme from *E.coli* 1100 strain resulted in about a fivefold increase in productivity as compared to that of late induction during the stationary phase. The highest values of specific activity and enzyme yield were obtained when 0.1mM IPTG was added to the medium at an optical density of 1.2 at 595nm. Therefore, the optimum time of induction for *Ecoli 294* (pPG430) was chosen to be the late exponential phase where the OD_{595} reached 1.2.

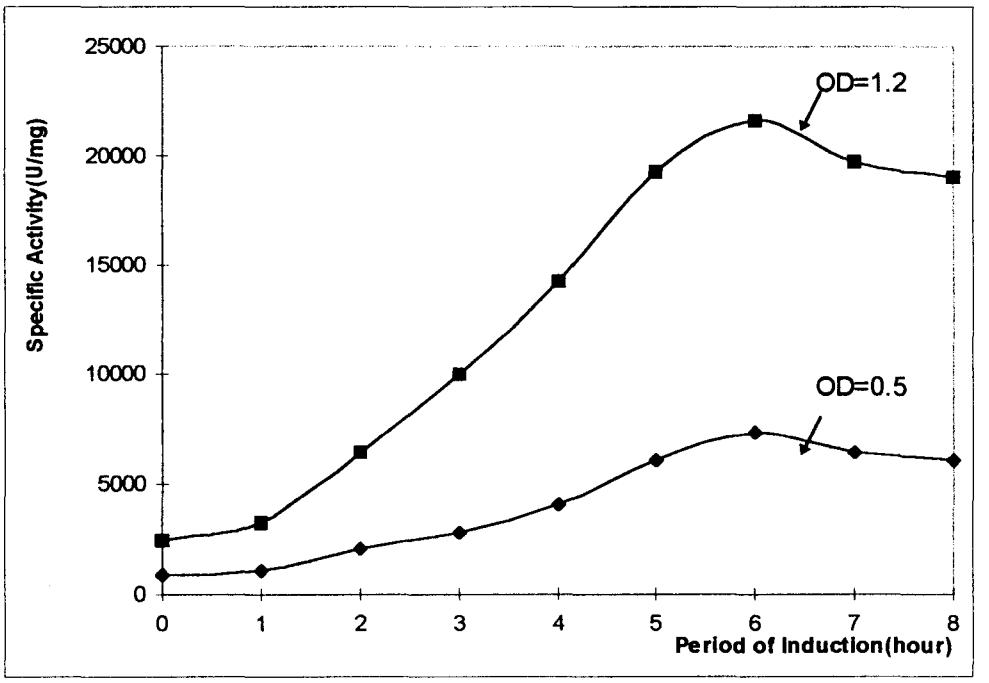


FIGURE 4.3. The effect of the period of induction on the *Eco*RI specific activity

After determining the optimum time and period of induction as an optical density of 1.2 at 595nm and 6 hours respectively, *the effect of the inducer (IPTG) concentration* on the efficiency of the induction was examined by using 0.1, 0.5 and 1.0mM IPTG concentrations. Three parallel shake-flask fermentations were performed, all cultures were grown until the OD_{595} reached 1.2 in the late exponential phase. 0.1, 0.5 and 1.0 mM IPTG was then added to each culture flask respectively, and the flasks were incubated for 6 hours. Cell extracts were separately assayed for *Eco*RI endonuclease activity. A control experiment was also performed in which the addition of IPTG was omitted. In the control experiment, the cells were also grown until the late exponential phase and then the cell extracts were analyzed for *Eco*RI activity. Table 4.2 shows that IPTG induction has

resulted in at least a 10 fold increase in the total activity and the yield of the enzyme. However, an increase in the IPTG concentration from 0.1 mM to 1.0 mM has not resulted in any significant improvement. Lee and Ramirez (1992) have also concluded that the protein level in *E.coli* cells increases rapidly following the induction of IPTG at a relatively lower IPTG concentration. At the higher IPTG concentrations, they have found almost constant productivity and concluded that IPTG has an inhibition effect on cell growth.

TABLE 4.1. The effect of the time of induction on the production of *Eco*RI endonuclease from *E.coli* 294(pPG430)

	OD_{595}	**C_{IPTG} (mM)**	**Volume (mL)**	**Total Protein (mg)**	**Total Activity (U)**	**Specific Activity (U/mg)**	**Yield (U/gcell)**
1	0.45	0.1	100	246	$1.5x10^6$	6098	$3.5 x10^5$
2	0.92	0.1	100	328	$4.0 x10^6$	12195	$9.0 x10^5$
3	1.2	0.1	100	416	$9.8 x10^6$	23558	$2.1 x10^6$
4	1.4	0.1	100	398	$7.5 x10^6$	18844	$1.7 x10^6$

TABLE 4.2. The effect of IPTG concentration on the production of *Eco*RI endonuclease from *E.coli* 294(pPG430)

	OD_{595}	**C_{IPTG} (mM)**	**Volume (mL)**	**Total Protein (mg)**	**Total Activity (U)**	**Specific Activity (U/mg)**	**Yield (U/gcell)**
1	1.2	0.0	100	202	$6.0x10^5$	2970	$1.5 x10^5$
2	1.2	0.1	100	403	$1.05 x10^7$	26054	$2.4 x10^6$
3	1.2	0.5	100	438	$1.2 x10^7$	27397	$2.9 x10^6$
4	1.2	1.0	100	452	1.3 x107	28761	$3.0 x10^6$

Since the cost of IPTG constitutes a major expense in the production, the lower IPTG concentration, 0.1 mM, was chosen as the optimum concentration for further experiments during which the main consideration was the optimum production of the restriction enzyme *Eco*RI .

4.2. Optimization of the Temperature-Shift Period for the Production of *Eco*RI *by E.coli M5248* (pSCC2)

In *E.coli M5248* strain containing pSCC2 plasmid, *Eco*RI structural genes for the restriction endonuclease and the modification methylase were oriented for expression under p_L promoter control. The initiation codon of the endonuclease gene is about 650 basepairs downstream from the p_L promoter (Figure 4.4). The synthesis of the *Eco*RI endonuclease is enhanced upon induction of p_L expression in strains producing a thermolabile λcI857 repressor. At 30-32°C, the repressor is active and the promoter is kept under repression; at elevated temperatures, productivity level increases due to the denaturation of the cI repressor (Cheng et al., 1984). The length of the temperature-shift period affects the product yield. However, it is also important to expose the culture to a higher induction temperature for a limited period for minimizing the negative effects of the thermal promotion on cellular metabolism (diPasquantonio et al., 1987)

In this part of the study, recombinant *E.coli* M5248 cells containing plasmid pSCC2 were thermally induced for different time periods varying from 1 to 8 hours, and the effect of temperature shift on product yield was investigated.

The growth behaviour of *E.coli M5248* containing the plasmid pSCC2 is shown in Figure 4.5. The exponential phase of the cell growth was completed in 12 hours after inoculation. A lag phase of around 3 hours was observed during the growth of the cells. Cells were always induced at an absorbance value of 1 at 590nm as suggested by Cheng et al (1984).

Eight parallel shake flask fermentations were performed in which recombinant cells were grown at 30-32°C in supplemented L-broth to an absorbance value of 1.0 at 590 nm. The culture temperature was then raised to 42°C by transferring all of the recombinant cultures to another shaker at 42°C. Each fermentation flask was held at 42°C for a defined period of time varying from 1 to 8 hours. A control experiment was also carried out with

the recombinant cell culture without exposing it to a temperature shift in order to follow the efficiency of the promoter.

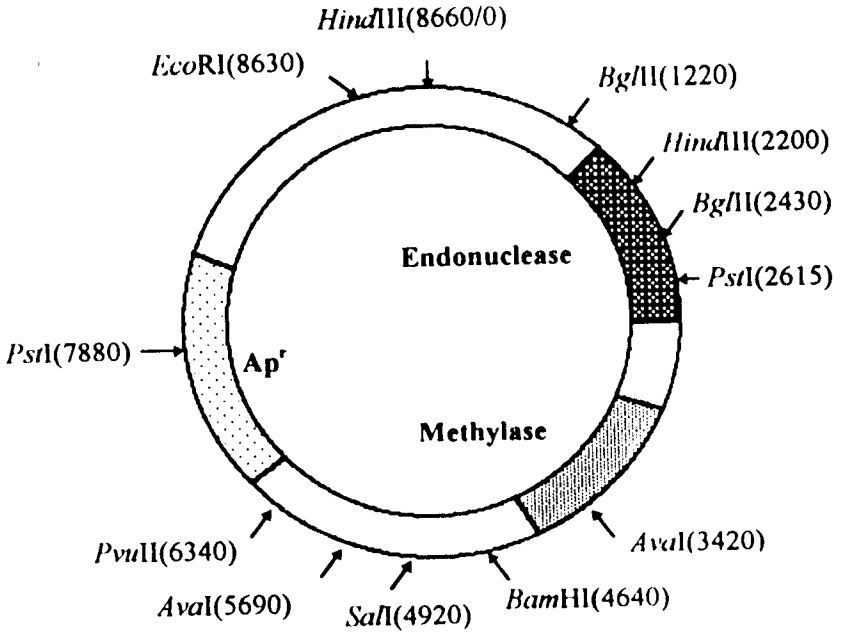


FIGURE 4.4. Structure of Plasmid pSCC2

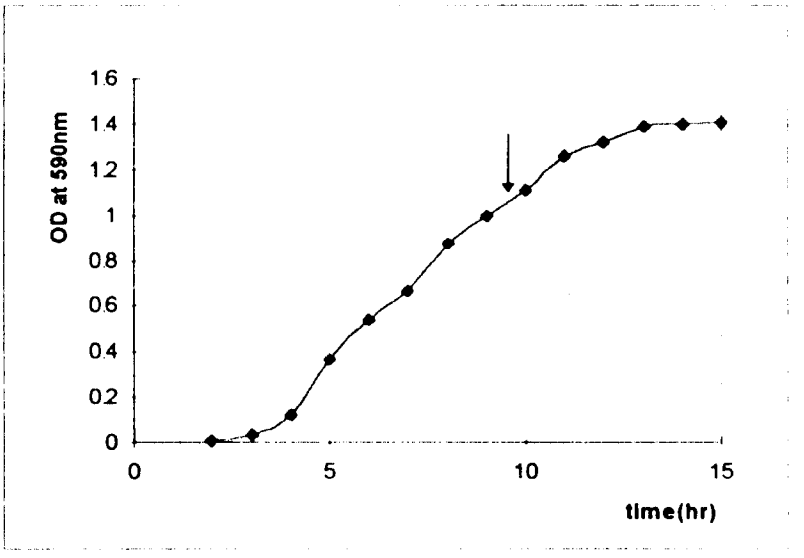


FIGURE 4.5. Growth curve for *E.coli M5248* cells (pSCC2) at 32°C (arrow indicates where the temperature shift from 30-32°C to 42°C was made).

The activities of the *Eco*RI endonuclease in the cell extracts of the recombinant culture exposed to different temperature-shift periods were compared to determine the effect of the temperature pulse on the *Eco*RI endonuclease productivity (Table 4.3). A gradual increase in the specific activity of *Eco*RI and in the protein content was observed up to 5 hours of T-shift period followed by a decrease. The proteins in the recombinant cultures exposed to longer temperature-shift periods could have been denatured, resulting in a decrease in the total protein amount and activity of the enzyme. Temperature-shift has resulted in at least 100 fold increase in the total activity and the yield of the enzyme. Total protein amount has also increased five fold at the optimum temperature period which was found to be five hours.

TABLE 4.3. The effect of temperature-shift on the *Eco*RI enzyme activity

T-shift Period (hour)	Total Protein (mg)	Total Activity (U)	Specific- Activity (U/mg)	Yield (U/g cells)
0	101	$5x10^3$	49.5	$1.31x10^3$
1	123	$1.25x10^4$	101.4	$3.3\ x10^3$
2	143	$1.5\ x10^4$	104.6	$4.06\ x10^3$
3	321	$1.5\ x10^5$	467.7	$3.6\ x10^4$
4	488	$2.5\ x10^5$	512.3	$6.2\ x10^4$
5	517	$5.0\ x10^5$	967.0	$1.2\ x10^5$
6	467	$4.0\ x10^5$	856.0	$9.3\ x10^4$
7	351	$1.0\ x10^5$	285.2	$2.3\ x10^4$
8	240	$7.5\ x10^4$	312.4	$1.6\ x10^4$

The variation of the yield and the specific activity of the *Eco*RI endonuclease with time is presented in Figure 4.6. Both the yield and EcoRI activity reach their highest values following an induction of 5 hours at 42°C. A slight decrease in the specific activity and the yield of the enzyme was observed for longer periods of induction.

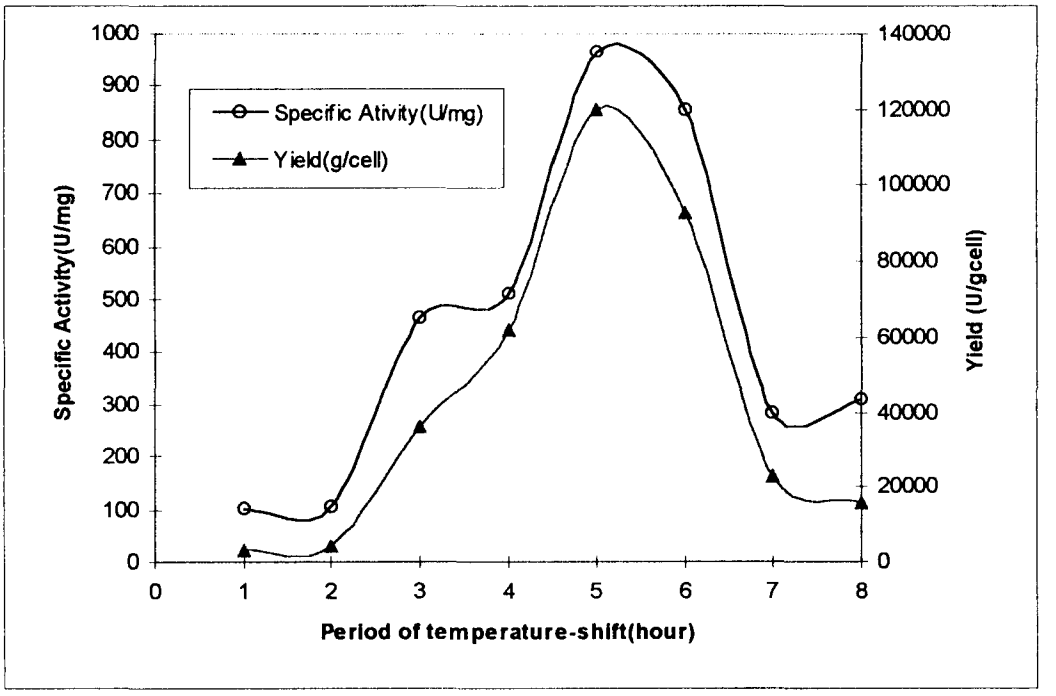


FIGURE 4.6. The effect of temperature-shift on the yield and the specific activity of *Eco*RI endonuclease (*E.coli* M5248).

The effect of variations in the length of the T-shift period on the plasmid stability was investigated by calculating the recombinant cell percentage in the fermentation flasks exposed to different T-shift periods. This percentage was defined as the ratio of colony numbers grown on an antibiotic containing plate to the total colony number observed in the absence of antibiotic. The percentage of the viable recombinant cells exposed to the different T-shift periods is presented in Figure 4.7. The number of plasmid containing cells was found to decrease as the temperature-shift periods increased (detailed explanation and related figures are given in Section 4.5.1). After five hours of temperature-shift, the number of plasmid bearing cells was reduced to less than 50%, resulting in a rather poor recombinant fermentation culture for the production of *Eco*RI endonuclease.

Although these results have indicated that a temperature shift period of 5 hours is the optimum for an efficient increase in the synthesis and recovery of EcoRI endonuclease, the negative effect of this period on the plasmid stability needs to be taken into consideration in further experiments.

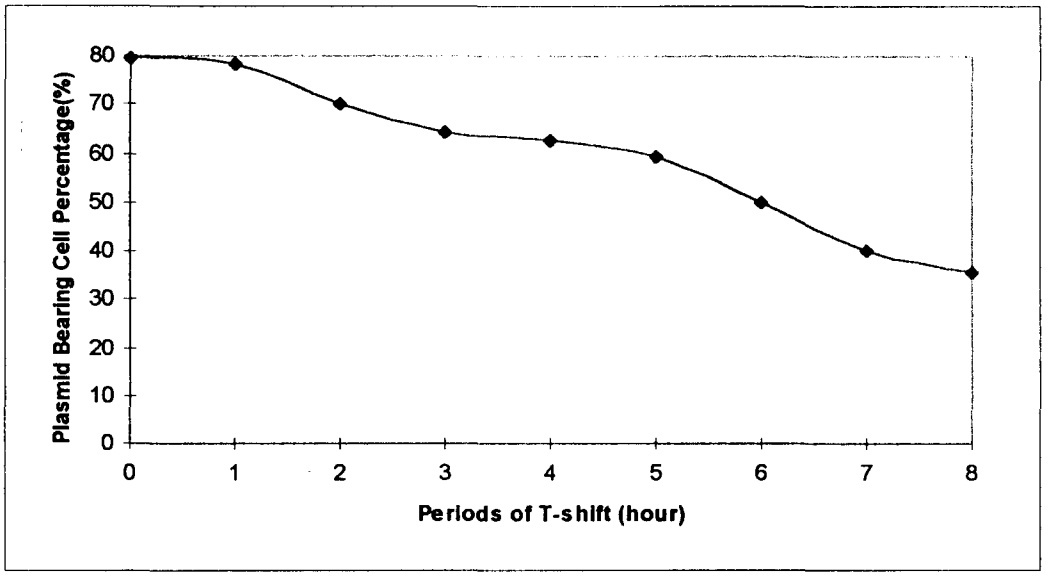


FIGURE 4.7. The effect of temperature shift on the recombinant plasmid stability of *E.coli* M5248.

4.3 Purification of EcoRI Restriction Endonuclease From Two Different Recombinant *E.coli* Strains

*Eco*RI endonuclease was purified both from the genetically modified *E.coli* 294 (carrying plasmid pPG430) and *E.coli* M5248 (carrying plasmid pSCC2) overproducing strains, and the enzyme yields were compared. In plasmid pPG430, genes encoding *Eco*RI endonuclease and methylase are oriented under the control of lac promoter, whereas in plasmid pSCC2, they are placed under the control of p_L. The purification of the enzyme from both strains was performed by using a modified version of the procedure developed for the purification of *Eco*RI endonuclease from an overproducing strain to minimize the formation of insoluble intracellular aggregates (Luke and Halford, 1985).

Cells that had been induced for the synthesis of *Eco*RI either by the addition of 0.1 mM IPTG at an optical density of 1.2 following a 6 hours induction period in the case of *E.coli* 294 (pPG430) or exposing the culture to a temperature-shift at 42°C at an optical

density of 1.0 for a period of 5 hours in the case of *E.coli* M5248 (pSCC2), were harvested at 9500 rpm for 15 minutes at 4°C using Sorvall GSA rotor. Collected cell pellets were then resuspended in ice-cold 0.9% NaCl and collected by centrifugation at the end of each fermentation. Approximately 4g/L of biomass was obtained from each strain

Cells were suspended in 25 mL of lysis buffer per gram of wet cells and disrupted by sonication. Cell disruption was followed spectrophotometrically over the different sonication periods tested for both recombinant *E.coli* strains (Figures 4.8- 4.9). The degree of distruption of the cell membrane was measured at 550nm, while the apperance of proteins in the cell extract was measured at 280nm. The optimum sonication period was chosen as 5 x 30 seconds for both strains, since protein content at longer sonication periods remained almost constant and the risk of denaturation of the enzyme increases at longer sonication periods.

The sonicated cell extract was dialyzed against Buffer A + 0.4 M NaCl to decrease the ionic stregth of the cell extract resulting from the composition of the lysis buffer and to avoid the aggregation of the protein. The cell debris was removed together with precipitated proteins by centrifugation at 9500rpm in a Sorvall SS-34 rotor for 15 minutes. The fraction at this step was named as Fraction I.

Phosphocellulose column was prepared and the column was equilibrated with Buffer A + 0.4M NaCl until the pH of the eluent was exactly the same as the washing buffer. Following the equilibration of the column, Fraction I was applied to the column as described in the Materials and Methods Section (3.2.7.(b)). The subsequent elution was carried out stepwise by Buffer A containing increasing concentrations of NaCl (from 0.4 to 1.0M). The fractions collected during elution were monitored spectrophotometrically at 280nm The elution profiles of the first phoshocellulose column are shown in Figures 4.10 and 4.11 for the two strains. Before the addition of a higher concentration of NaCl, the absorbance values of the fractions were allowed to reach the baseline value. The fractions that have high protein content were assayed for EcoRI endonuclease activity. Active enzyme was found to be eluted by 0.6M NaCl concentration. The active fractions were pooled for the following chromatographic step and named as Fraction II.

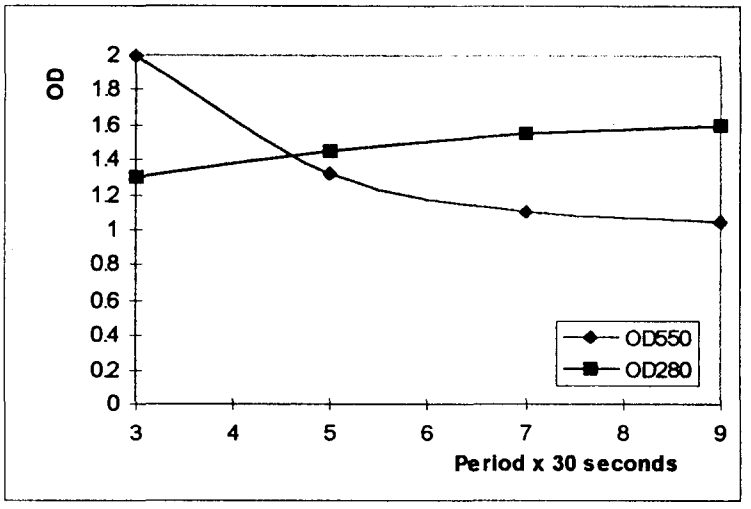


FIGURE4.8. Determination of the sonication period for *E.coli 294* (pPG430) cells

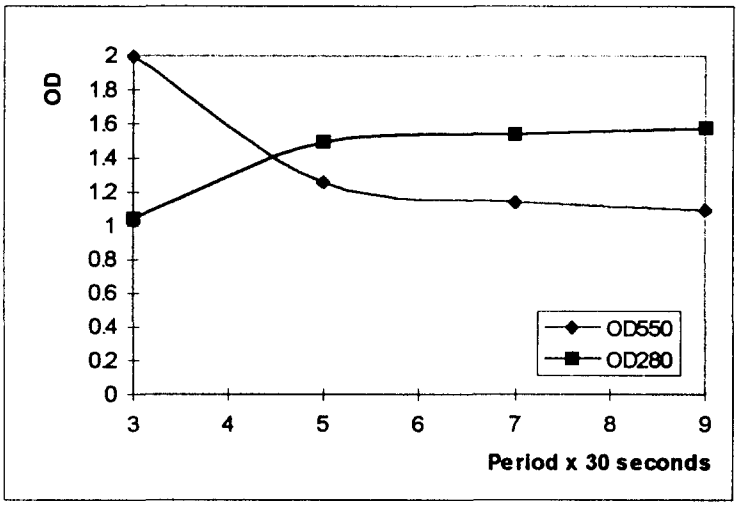


FIGURE 4.9. Determination of the sonication period for *E.coli M5248* (pSCC2) cells

Fraction II of each strain was mixed with the slurry of hydroxyapatite which was pre-equilibrated with Buffer A + 0.6 M NaCl and centrifuged. The pellets were washed with increasing concentrations of K-phosphate ranging from 0.1M to 0.6M in Buffer A containing 0.6M NaCl. Active enzyme was found to be eluted by 0.4M K-phosphate concentration and was named as Fraction III. Batchwise hydroxyapatite chromatography has resulted in 71 and 96% recovery of *Eco*RI endonuclease from the two different *E.coli* strains, *E.coli M5248* and *E.coli 294*, respectively (Table 4.4).

Fraction III of each strain was diluted four times with Buffer A and applied to the second phosphocellulose column which was eluted in a manner similar to the first one. In the case of strain M5248 , the application of a second phosphocellulose column could increase the specific activity of the *Eco*RI three-fold by eliminating contaminating proteins (Table 4.4). The elution profile of the second phosphocellulose column for *E.coli* M5248 strain is shown in Figure 4.12.

The application of the second phosphocellulose column has not resulted in an increase in the specific activity of *Eco*RI in the case of *Ecoli* 294 strain but caused a significant decrease in the recovery of the enzyme (Table 4.4). As it could not improve the purification any more, the application of the second phosphocellulose column could be omitted in the case of *E.coli* 294 (pPG430). This observation allowed the development of a simple two step procedure consisting of only two chromatographic steps for the preparation of pure *Eco*RI for commercial use from this genetically engineered overproducing strain.

Table 4.4 summarizes the purification *Eco*RI endonuclease from *E.coli* 294 (pPG430) and *E.coli* M5248 (pSCC2), respectively. An approximately 8-fold purification of proteins was achieved from *E.coli* 294 (pPG430) strain after the first phosphocellulose chromatography whereas this value was 14 fold for *E.coli* M5248 (pSCC2) strain. The application of hydroxyapatite chromatography following the first phosphocellulose chromatography resulted in about 3 fold purification of proteins accompanied by approximately 1.1 fold decrease in the total activity of both recombinant strains.

The activity measured in the clarified cell extract of the *Eco*RI endonuclease isolated from *E.coli* 294 (pPG430) was found to be 3.43×10^6 units whereas it was 2.13×10^5 units per gram wet cells in the case of *E.coli* M5248 (pSCC2). The existence of this 16 fold difference in the clarified cell extracts have clearly indicated that *E.coli* 294 with pPG430 is a better source for the production of *Eco*RI endonuclease. The application of a 2-step protocol for *E.coli* 294 (pPG430) and a 3-step protocol for *E.coli* M5248 (pSCC2) resulted in a production yield of 3.3×10^6 units and 1.3×10^5 units of *Eco*RI endonuclease per gram wet cells with a recovery of 96% and 61% respectively. 3.1 and 1.2 mg final product per

gram wet cells were obtained with specific activities of $1x10^6$ U/mg and $1x10^5$ U/mg for *E.coli* 294 (pPG430) and *E.coli* M5248 (pSCC2), respectively.

TABLE4.4. Purification of *Eco*RI endonuclease from the recombinant strains

E.coli 294 **(pPG430) Strain (4.375gcells)**						
Fractions	**Steps**	**Volume (ml)**	**Total Protein (mg)**	**Total Activity (U)**	**Specific Activity (U/mg)**	**Recovery (%)**
I	Supernatant	100	357	$1.5x10^7$	42 031	100
II	Phosphocellulose	84	44	$1.5x10^7$	336 000	98
III	Hyroxyapatite	48	14	$1.4x10^7$	1 045 372	96
IV	Phosphocellulose	55	9	$9.6x10^6$	1 066 718	64
E.coli ***M5248*** **(pPG430) Strain (4.923gcells)**						
I	Supernatant	150	932	$1.05x10^6$	1127	100
II	Phosphocellulose	80	64	$8.9x10^5$	13 906	85
III	Hyroxyapatite	48	24	$7.5x10^5$	30992	71.4
IV	Phosphocellulose	64	6	$6.4x10^5$	104 918	60.9

Cheng et al (1984) have purified *Eco*RI endonuclease from *E.coli* M5248 (pSCC2) by using a different protocol consisting of streptomycin and ammonium sulfate fractionations followed by phosphocellulose and hydroxyapatite chromatographies respectively. They have obtained 500mg of enzyme per kg cell paste with a recovery of 47%. The specific activity of the *Eco*RI endonuclease was reported to be $4.5x10^4$ U/mg. Therefore it has been calculated that approximately $2.25x10^4$ U enzyme per gram cell could be obtained by Cheng et al (1984). Almost a tenfold increase in the yield of the *Eco*RI endonuclease was obtained in the present study by the application of a new purification scheme ($6.4x10^5$ U enzyme obtained from 4.923 gcells).

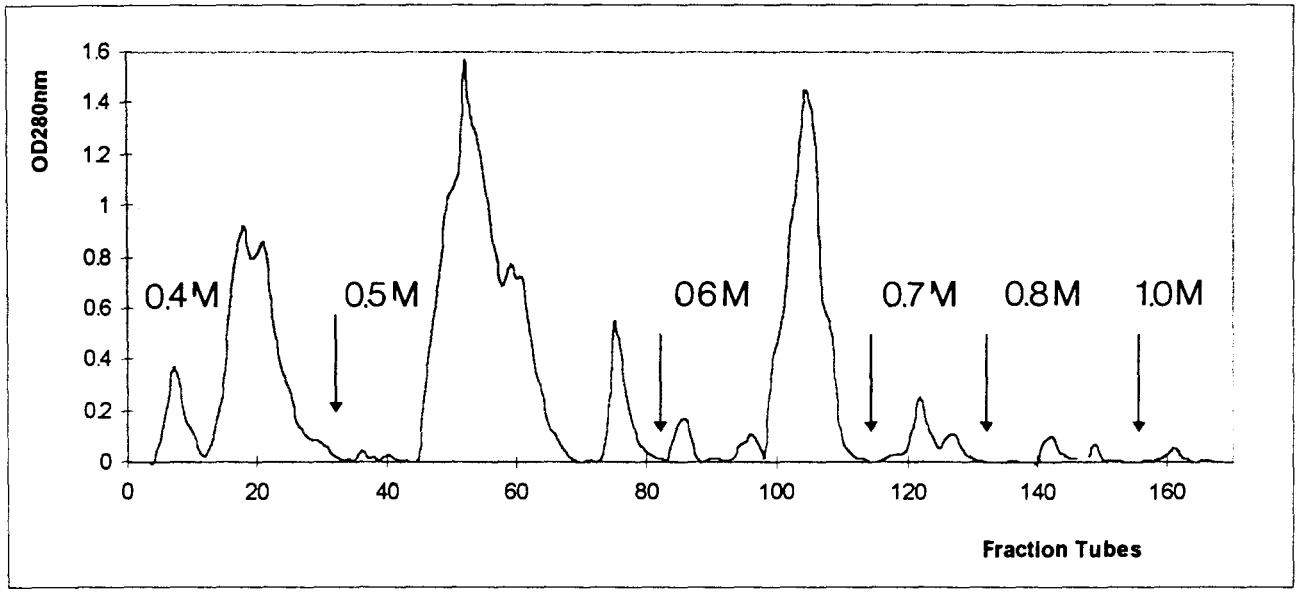


FIGURE 4.10. Elution profile of the first phosphocellulose column chromatography for *E.coli* 294 strain. (Arrows indicate the points where the buffer has been changed)

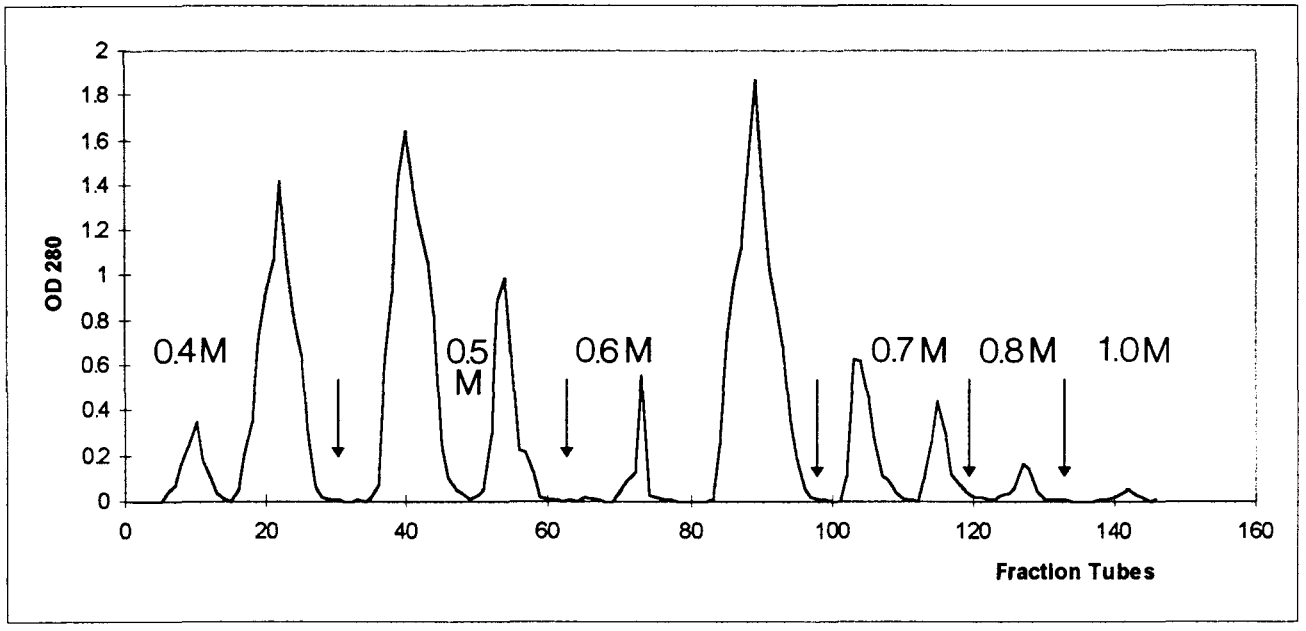


FIGURE4.11. Elution profile of the first phosphocellulose column chromatography for *E.coli* M5248 strain. (Arrows indicate the points where the buffer has been changed)

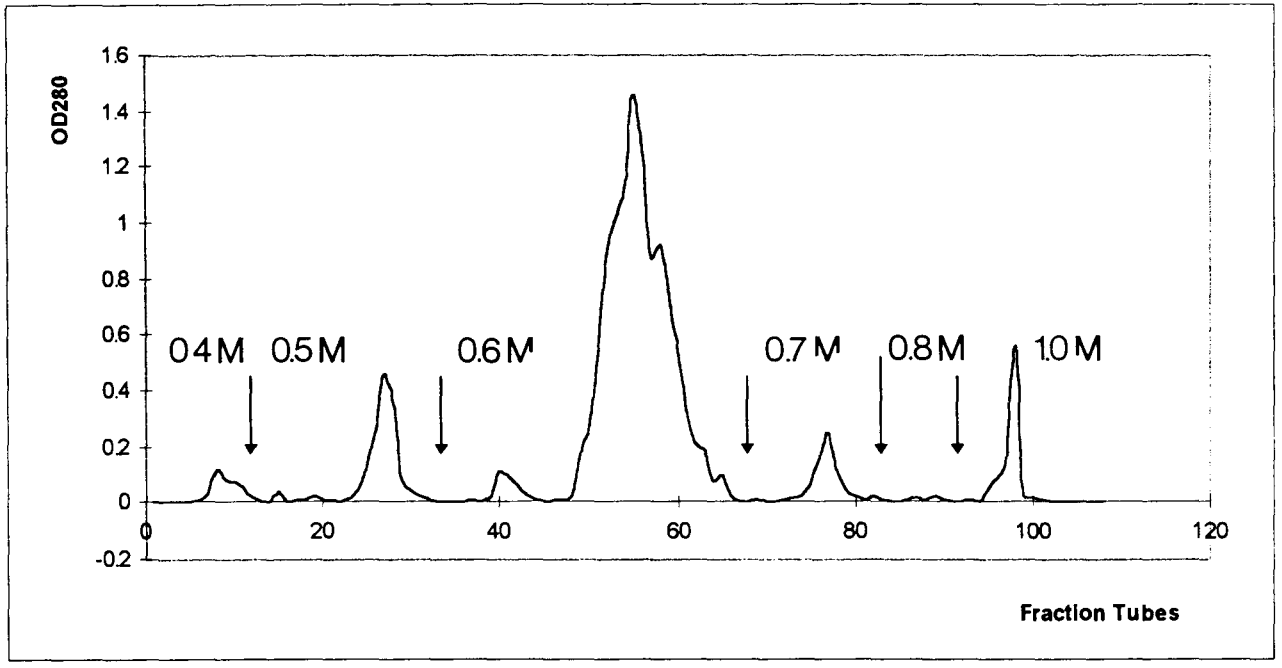


FIGURE 4.12. Elution profile of the second phosphocellulose column chromatography for *E.coli* M5248strain. (Arrows indicate the points where the buffer has been changed)

However, the yield and the specific activity of the enzyme produced by *E.coli* M5248 (pSCC2) remained lower compared to that obtained by Luke and Halford (1985). These investigators have used a different overproducing construct in which the gene encoding *Eco*RI was placed under the control of the same p_L promoter, and the genes carrying the *Eco*RI methylase and cI-coded temperature sensitive repressor were on compatible separate plasmids. These substantial differences in yield and specific activity of the enzyme between the two strains used by Cheng et al. (1984) and Luke and Halford (1985) may possibly be due to the lower expression of the strain M5248 (pSCC2). This lower expression may result either from the distance between the p_L and the gene for *Eco*RI endonuclease or from the simultaneous placement of *Eco*RI and methylase genes on the same plasmid. The nature of the host cells and plasmid copy number may be other important factors that lead to a lower enzyme recovery in the case of *E.coli* M5248 (pSCC2).

The higher productivity of *Eco*RI endonuclease from p_L promoter reported by Luke and Halford (1985) for their overproducing construct has indicated that lower product yield (U/g cell) obtained from *E.coli* M5248 compared to *E.coli* 294 strain may not only be explained by the presence of two different promoters. In order to get a better understanding of the productivity differences, the plasmid stabilities of these two constructs in which the gene encoding *Eco*RI endonuclease and methylase are on a single plasmid were investigated and compared (detailed explanation is given in Section 4.4).

The highest specific activity reported for *Eco*RI endonuclease purified from overproducing strains until now, was the one reported by Luke and Halford with a value of 2.18×10^7 U/mg protein. A specific activity of 1.07×10^6 U/mg protein was reached in the present study by using the recombinant system of *E.coli 294*(pPG430).

4.3.1. Purity of *Eco*RI Endonuclease

The final enzyme preparations were subjected to electrophoretic analysis under denaturing conditions. Electrophoretic analysis by SDS-PAGE of enzyme preparations obtained from these two different overproducing strains have shown an identical pattern with that of the commercial *Eco*RI (Figure 4.13). A single protein band referring to the dimer form of the enzyme was observed for both *Eco*RI preparations. The complete dissociation of the protein was not achieved under the denaturing conditions reported. In the literature, it is frequently stated that *Eco*RI endonuclease was observed in dimer or tetramer form even under denaturing conditions (Bingham et al., 1977; Rubin and Modrich, 1980).

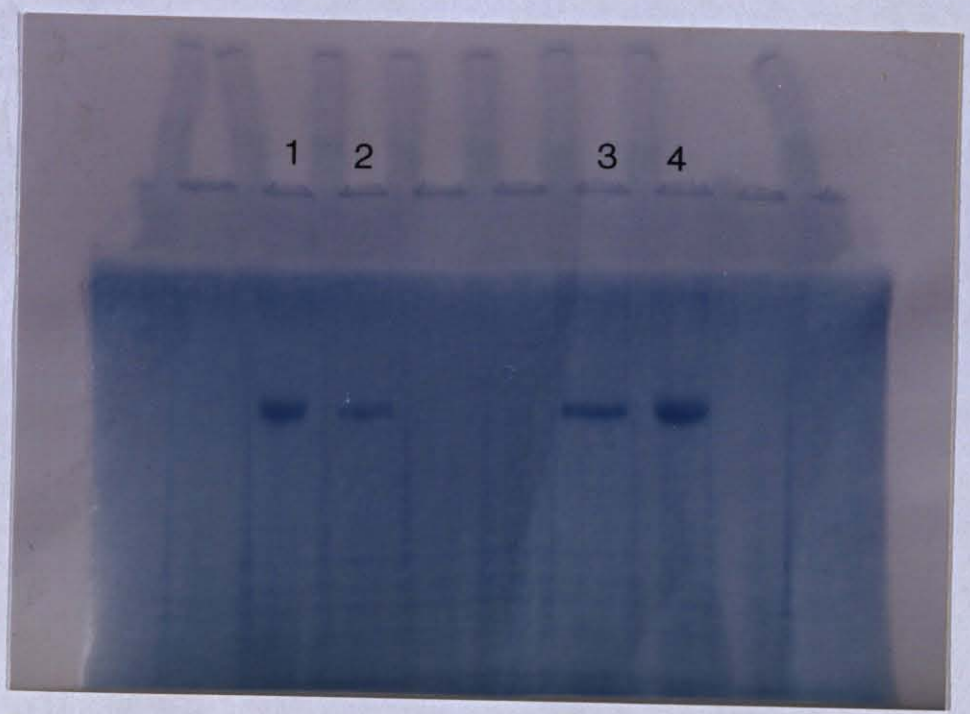


FIGURE 4.13. SDS-PAGE Analysis of the Protein Content.
Lane 2: *Eco*RI isolated from *E.coli 294* (pPG430) strain,
Lane 5: *Eco*RI isolated from *E.coli M5248* (pSCC2) strain.
Lane 1 & 6: Commercial *Eco*RI

4.3.2.The Quality Control Tests

4.3.2.1 Overdigestion Assay. An overdigestion quality test indicates the absence of endo- and exonucleases in the final enzyme preparations. Each final preparation of *Eco*RI was tested for contamination by other endodeoxyribonucleases capable of digesting DNA either at random sites or specific sites. For this purpose, one microgram of λDNA was digested with 20 units of the enzyme for 5 hours at an appropriate temperature. This represents 100 fold excess digestion as compared to 1 unit determination reaction at the end of 1 hour. Digestions were investigated on 0.8% agarose gel by visualizing under UV-light, in the presence of EtBr and photographed (Figure 4.14).

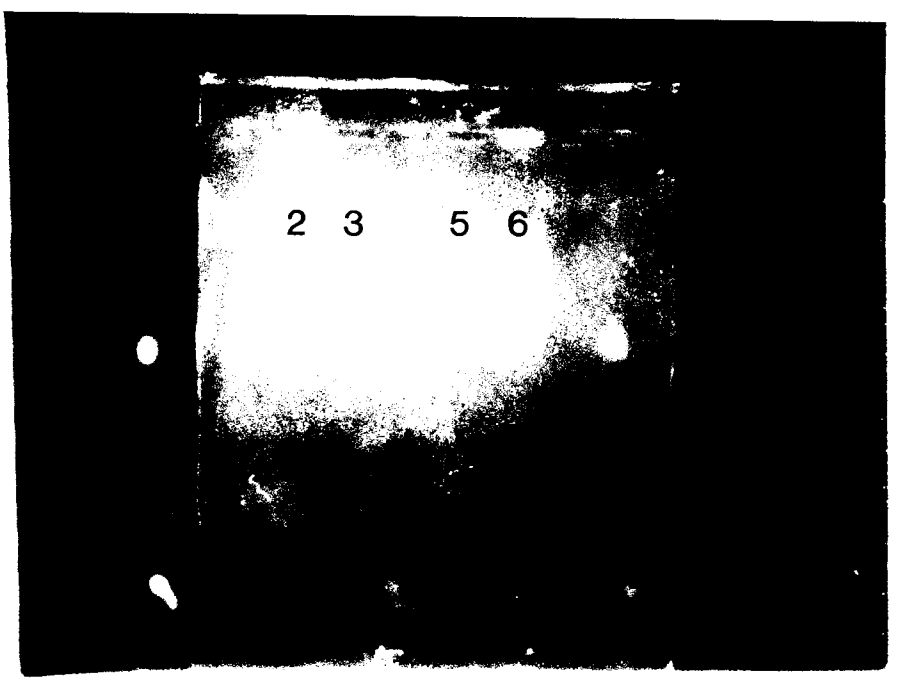


FIGURE 4.14. Overdigestion Assay of Purified *Eco*RI endonuclease.

Lane 2: *Eco*RI isolated from *E.coli 294* strain

Lane 3: Commercial *Eco*RI

Lane 5: *Eco*RI isolated from *E.coli M5248* strain

Lane 6: Commercial *Eco*RI.

The production of expected fragments at the end of the overdigestion assay have indicated that endo- and exonucleases were absent in the final enzyme prepartions purified from each recombinant strain.

4.3.2.2 Cut-Ligate-Recut Assay. The same *Eco*RI preparations were also tested for the presence of contaminants that would inhibit ligation or degrade termini. After performing the initial cleavage with *Eco*RI isolated in this study, DNA was extracted by phenol and chloroform and precipitated with ethanol. T4 ligase was used to ligate the fragments obtained from the initial cleavage. Ligation was performed at 16°C for 4 hours under the conditions described by the manufacturer (Mehra et al., 1993; Luke and Halford, 1985). T4 ligase was inactivated by heating for 15 minutes at 65°C. Ligated fragments were recut by using the same enzyme preparations.

In Figure 4.15, expected fragments were observed at the end of the initial cleavage of λDNA with *Eco*RI endonuclease for both strains. Ligated fragments resulted in the formation of the initial λDNA observed as a single band. Digestion of the product formed gave the expected fragments (Figure 4.15). The analysis of the fragments obtained after ligation and recut has shown that the final enzyme preparations were found to be free of contaminants that would inhibit ligation or degrade termini on 0.8% agarose gel.

These results have clearly shown that the preparations of the *Eco*RI endonuclease purified from each of the overproducing strains of *E.coli* in this study are suitable for use in molecular biology.

FIGURE4.15. Cut-Ligate-Recut Test of Purified *Eco*RI Endonuclease.

Lane 1: Initial cleavage of λDNA by *Eco*RI isolated from *E.coli294* (pPG430) strain

Lane 2: Initial cleavage of λDNA by *Eco*RI isolated from *E.coli M5248* (pSCC2) strain

Lane 4: Ligation of *Eco*RI isolated from *E.coli 294* (pPG430) strain

Lane 5: Ligation of *Eco*RI isolated from *E.coli M5248* (pSCC2) strain

Lane 7: Recut of *Eco*RI isolated from *E.coli M5248* (pSCC2) strain.

Lane 8: Recut of *Eco*RI isolated from *E.coli 294* (pPG430) strain

4.4 Investigation of the Plasmid Stabilities of *E.coli* 294 (pPG430) and *E.coli* M5248 (pSCC2)

Plasmid stability is an essential prerequisite for a successful production of foreign proteins by recombinant cells. It is often defined as the ability of transformed cells to maintain their plasmid unchanged during their growth. When using recombinant microorganisms, the objective is to be sure of their unchanged plasmid content over the cell growth period. In this part of the study, the stability of the two plasmids, pSCC2 and pPG430, were investigated in *E.coli strains M5248* and *294* respectively.

In order to examine the effect of the presence of the recombinant vector on the growth rate of the host cell, growth chracteristics of the recombinant *E.coli M5248* and *E.coli 294* cells were determined together with their respective native host cells. Batch fermentations were performed under the conditions described in the Materials and Methods, Section (3.3).

The effect of temperature upshift on the stability of the plasmid pSCC2, and the performance of cells exposed to elevated temperature after a subsequent temperature downshift, were investigated in the case of *E.coli* M5248.

The plasmid stability of pPG430 in the presence of selective and non selective media was investigated in the case of *E.coli* 294. Both plasmids have also been evaluated in terms of their stability after several successive generations in order to investigate the necessity to use freshly transformed cells for each fermentation.

4.4.1 Stability of Plasmid pSCC2 in *E.coli M5248* Cells

4.4.1.1 The Effect of Temperature-Shift on Plasmid Stability. In the present section, the main aim was to determine the effect of the temperature-shift on the recombinant cell viability and consequently on the *Eco*RI enzyme productivity. Shake flasks containing LB medium of 495 ml were inoculated with 5 ml of the overnight culture. *E.coli M5248* (pSCC2) cells were grown at 30-32°C in supplemented L-broth to an absorbance value of 1.0 at 590 nm. The culture temperature was then raised to 42°C for specified periods of time varying between 1-8 hours. Samples were taken from the culture broth every hour for the measurement of the absorbance and also for determination of the viable cell number, specific activity of *Eco*RI and the corresponding dry weight of the cells as described in Section 3.3.1. The percentage of the recombinant cell population was found by the replica-plate method.

Figures 4.16-4.23 show the variation in the biomass concentrations of recombinant cells without T-shift, recombinant cells and plasmid-free cells subjected to T-shift for different periods changing from 1 to 8 hours. The experiment was stopped after the T-shift period. The experiments were duplicated in all cases. It is observed from these Figures that the cultures immediately entered stationary phase after the T-shift. When calculating the μ_{max} values in each case, time intervals were chosen only in the exponential phase, and the accelaration phase following the lag phase was not included (Table 4.5).

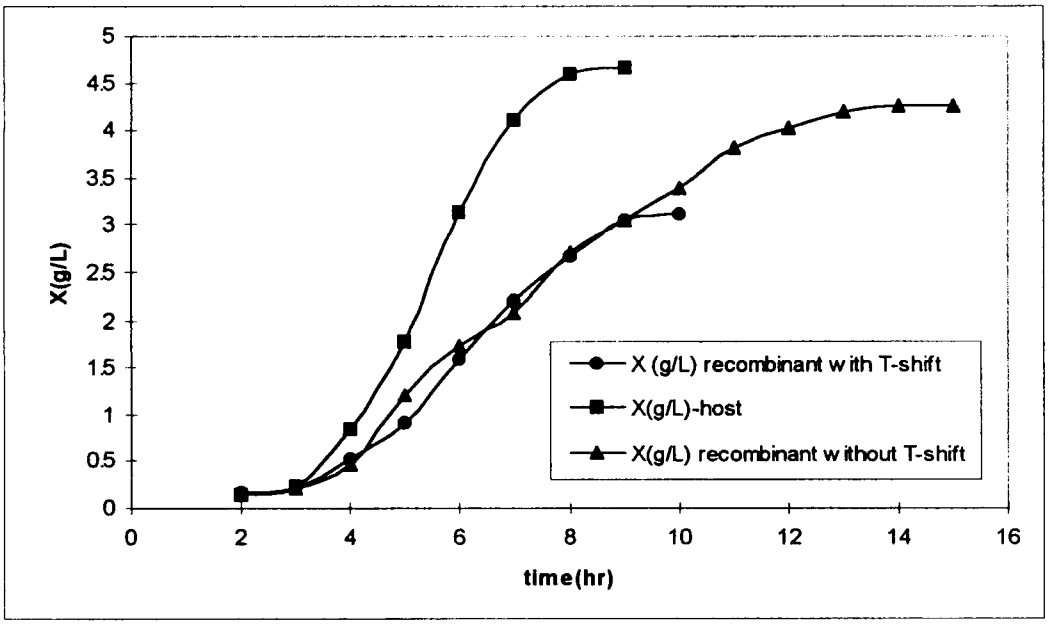


FIGURE 4.16. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 1 hour temperature-shift

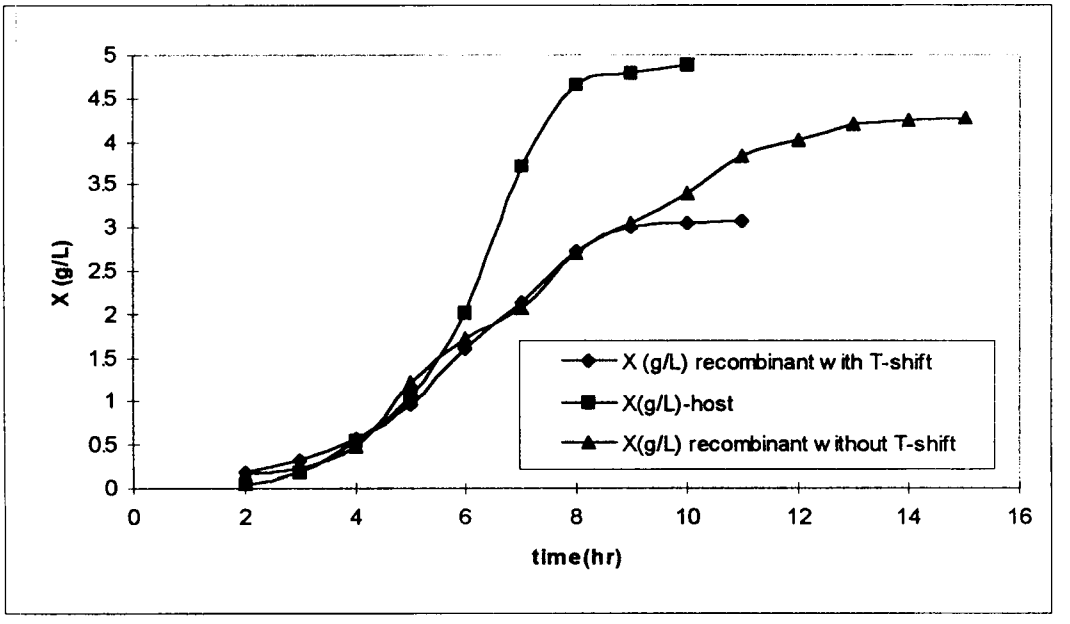


FIGURE 4.17. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 2 hour temperature-shift

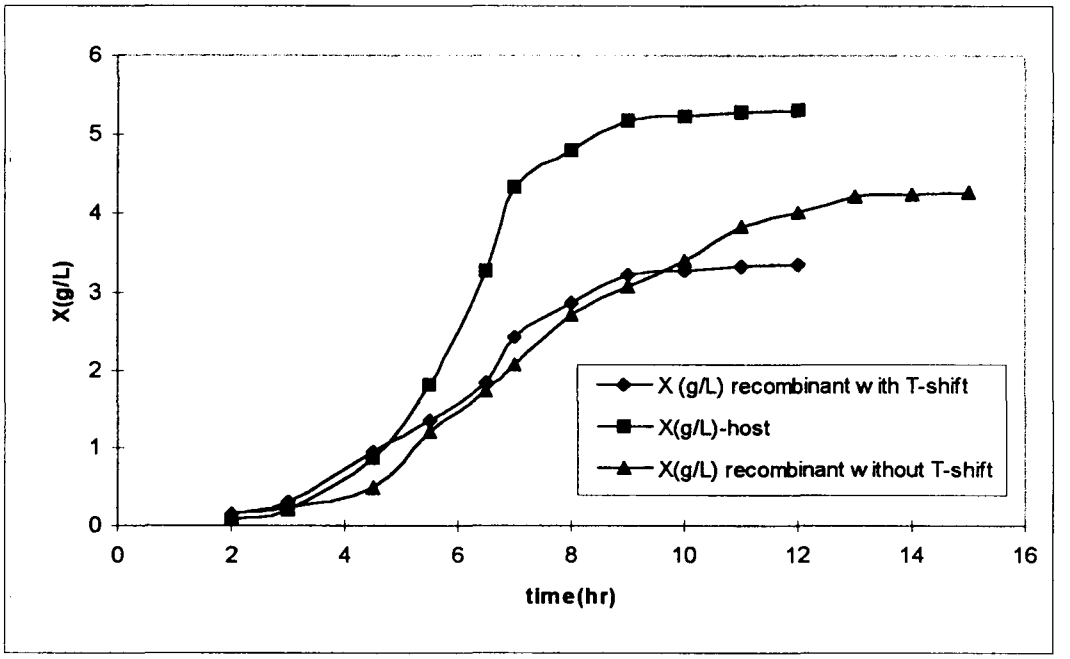


FIGURE 4.18. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 3 hour temperature-shift

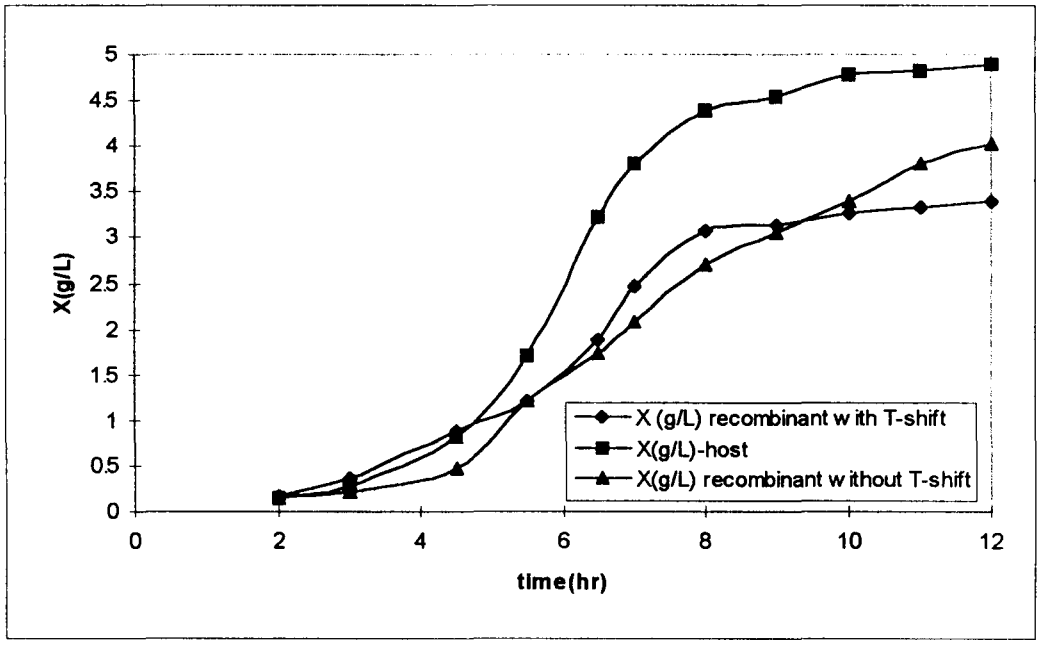


FIGURE 4.19. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 4 hour temperature-shift

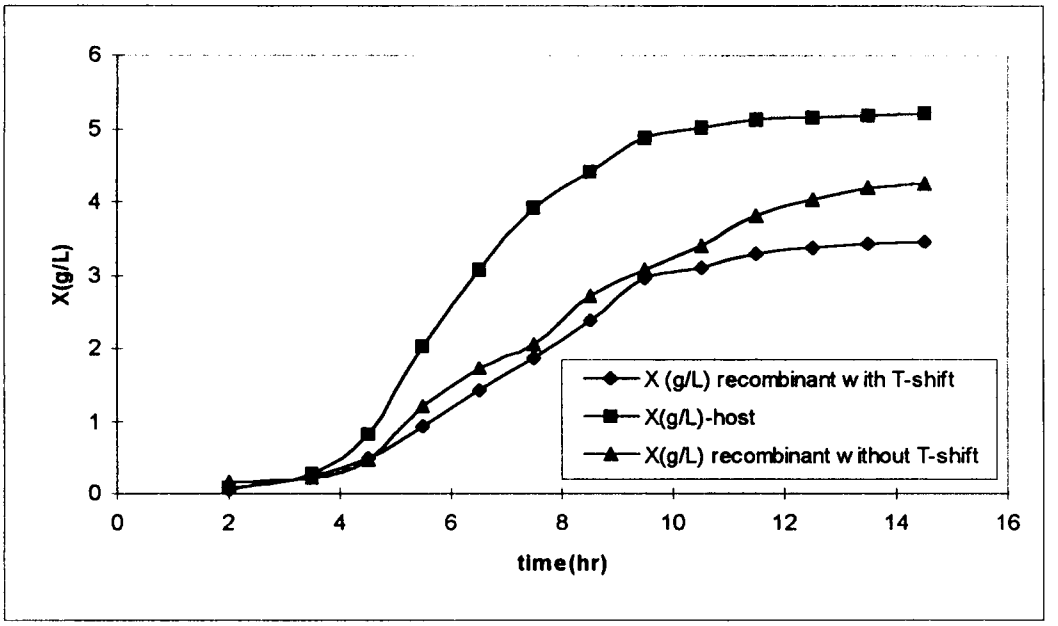


FIGURE 4.20. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 5 hour temperature-shift

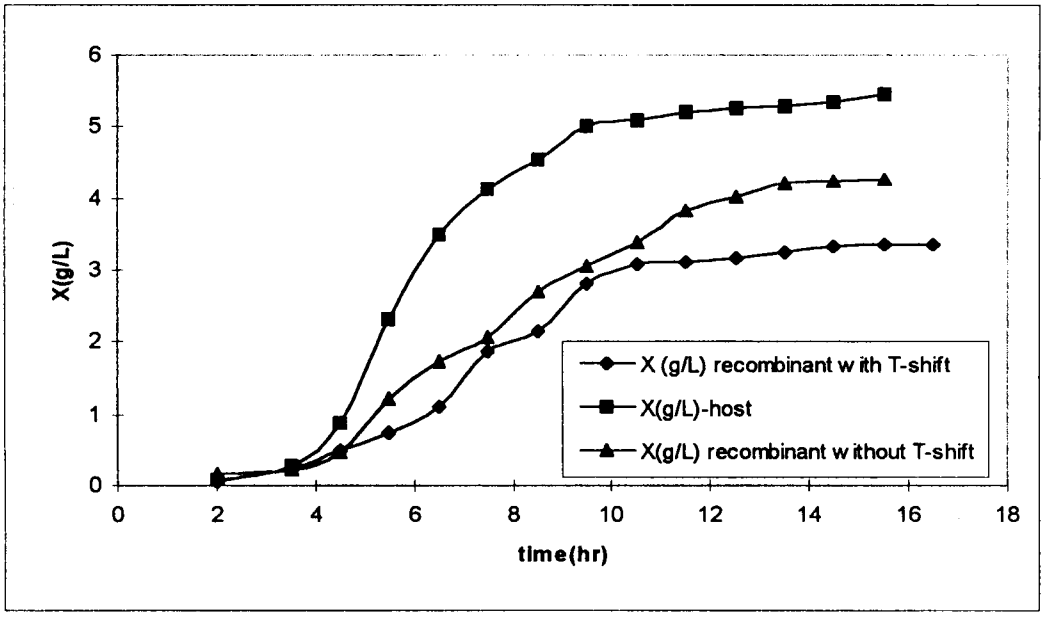


FIGURE 4.21. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 6 hour temperature-shift

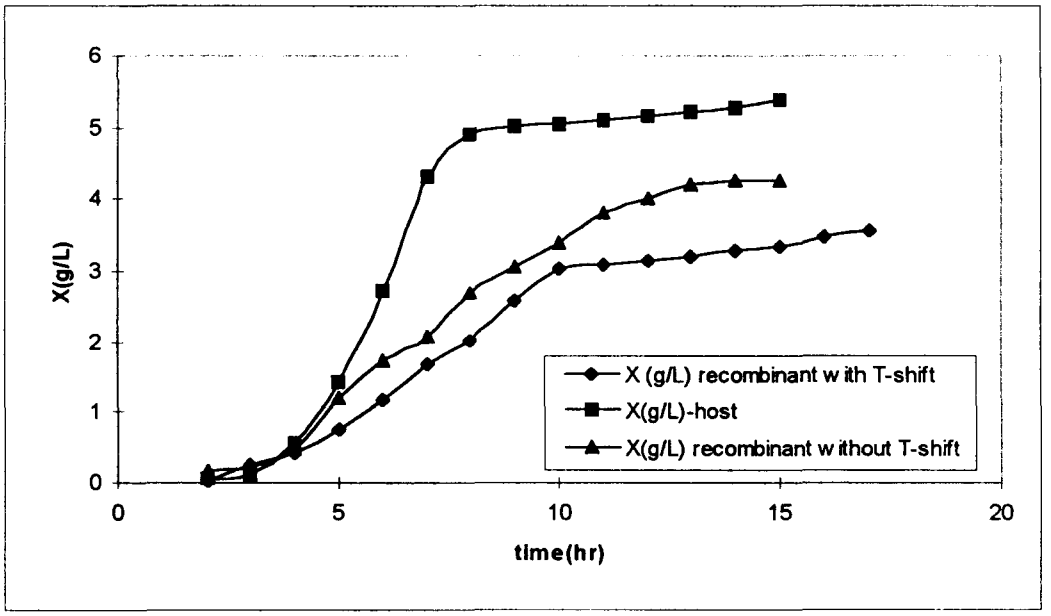


FIGURE 4.22. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 7 hour temperature-shift

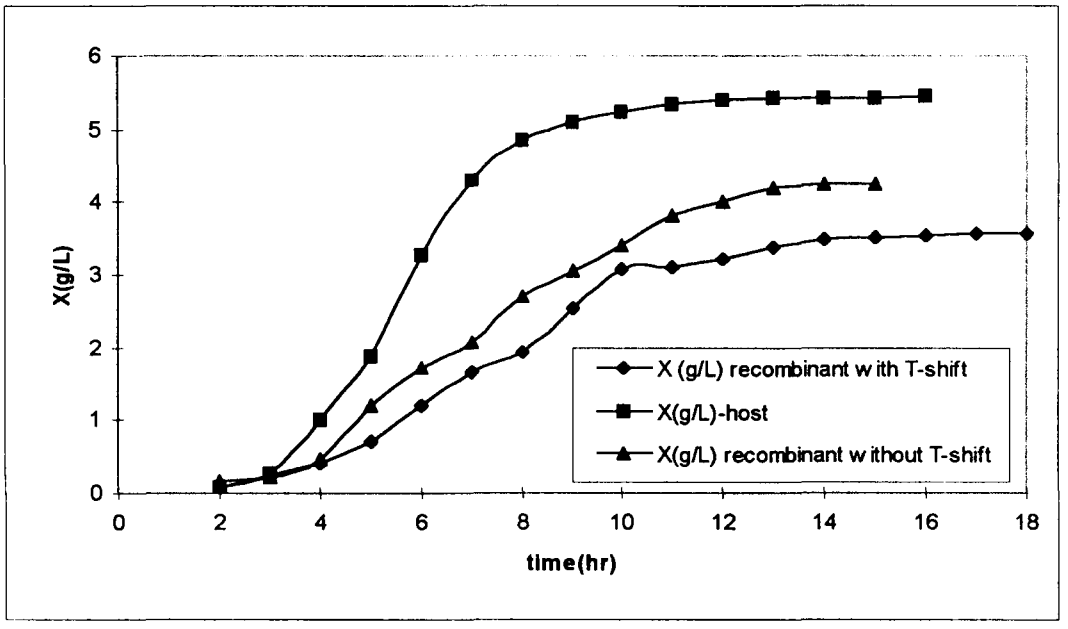


FIGURE 4.23. Growth curve of the host and the recombinant cultures of *E.coli* M5248 exposed to an 8 hour temperature-shift

The maximum specific growth rates calculated from growth curves in Figures 4.16-4.23 for 8 different temperature shift periods are given in Table 4.5. In the same table, the growth ratios calculated from the maximum specific growth rates of the plasmid harboring cells to that of the host cells are also given. Maximum specific growth rates determined from the exponential phase for the recombinant and the host cells are consistent with each other since the growth ratios indicated in the table are constant at 0.6 ± 0.04. As expected, the growth ratios were lower than one, indicating higher biomass production rates of biomass for host cells. There is a large difference between the growth rates of the recombinant and the host cells. One possible conclusion that can be drawn from these results is that plasmid instability may be due to growth rate difference rather than segregational instability in the case of *E.coli* M5248(pSCC2). It has been shown that growth rate difference and the segregational instability (i.e. the rate at which plasmid-free cells are generated from plasmid-bearing cells) are the two important physiological parameters involved in the generation of plasmid-free cells (Cooper et al.,1987).

TABLE 4.5. Maximum Specific Growth Rates for Recombinant and Host Cells and Growth Ratios of *E.coli* M5248

T-shift Periods (hr)	**μ^+_{max} ($\pm$ 4%) (recombinant cells) (hr^{-1})**	**μ^-_{max} ($\pm$ 3%) (host cells) (hr^{-1})**	**Growth Ratio ($\mu^+_{max} / \mu^-_{max}$)**	**Growth Rate Differences ($\mu^-_{max} - \mu^+_{max}$) ($hr^{-1}$)**
1	0.411	0.698	0.589	0.287
2	0.395	0.642	0.615	0.247
3	0.362	0.638	0.567	0.272
4	0.403	0.608	0.662	0.205
5	0.391	0.663	0.589	0.272
6	0.402	0.680	0.591	0.278
7	0.397	0.669	0.593	0.272
8	0.392	0.676	0.580	0.284

The change in the number of viable cells and in the percentage of plasmid containing cells with respect to time in batch cultures exposed to T-shift periods between 1-8 hours are given in Figure 4.24 and 4.25 respectively. In duplicate runs for different temperature shift periods, similar values were obtained for the percentages of plasmid-bearing cells and the cell viability. Cell counts indicate that the cells lose viability after they are exposed to elevated temperature, and as the length of the temperature shift period increases cell viability decreases gradually. The plasmid containing cell percentage also decreases after exposing to elevated temperature. There is a steady decrease in the plasmid containing cell percentage after the T-shift (See also Figure 4.7). After five hours, the plasmid containing cell percentage drops below 50%, and it reaches around 35% after the eight hour period. This result confirms that T-shift periods longer than five hours result in very poor fermentation in terms of the recombinant product.

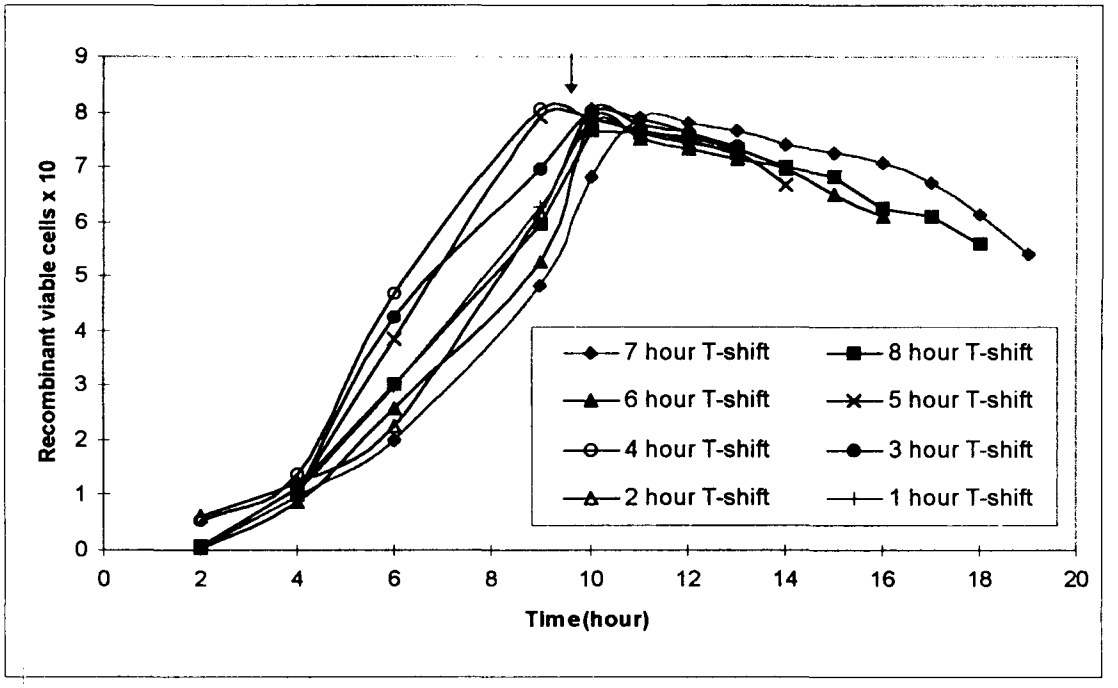


FIGURE 4.24. The change in recombinant viable cell counts with respect to time in batch cultures exposed to different temperature-shift periods between 1-8 hours (arrow indicates when T-shift is made)

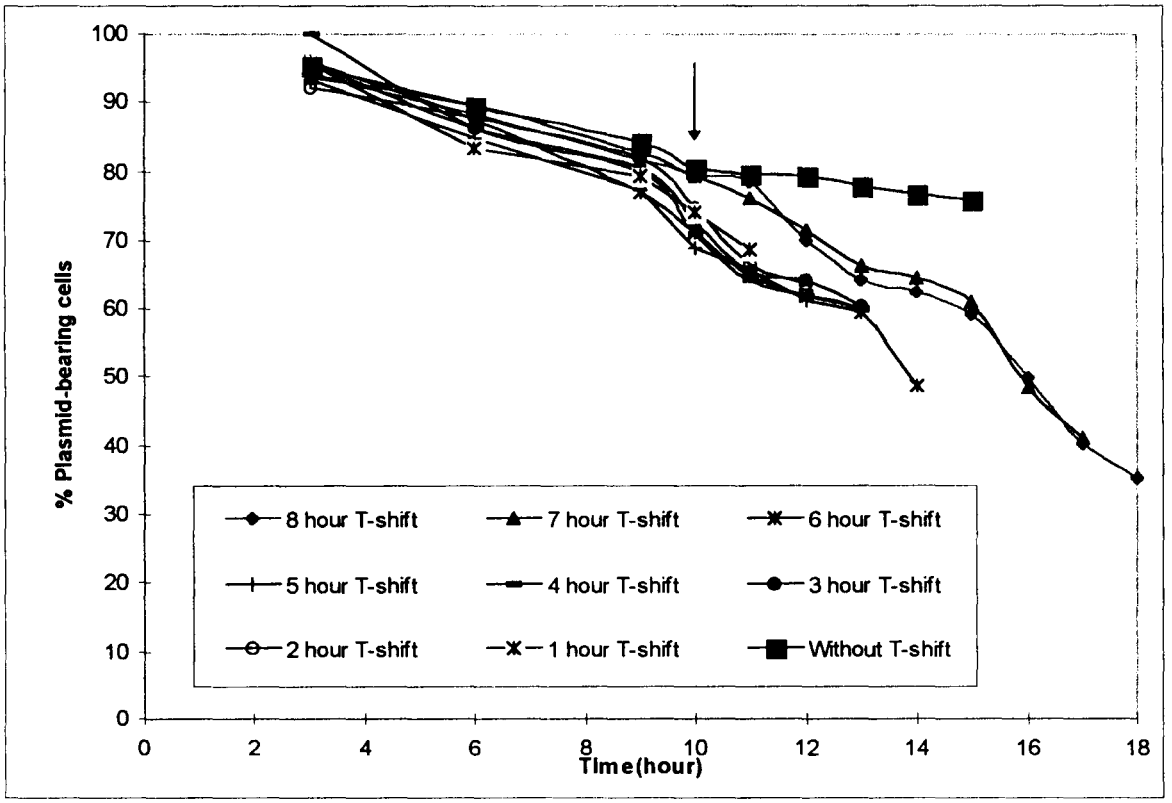


FIGURE 4.25. The change in the percentage of plasmid containing cells with respect to time in batch cultures exposed to different temperature shift periods between 1-8 hours(arrow indicates when T-shift is made)

In Figure 4.26, the change in specific growth rates, μ, and the *Eco*RI endonuclease production for the temperature shift period of 8 hours are plotted. The T-shift is made in the late exponential phase as cell growth starts to slow down. The μ values correspond to point rates along the growth curve after the T-shift and were calculated at each hour using the relation (1/ X) ΔX/Δt. This plot demonstrates that the effect of T-shift is to decrase recombinant cell growth. The decrease observed in the first two hours is rapid followed by an attempt for adaptation and eventually zero growth. An increase was observed in the production of the *Eco*RI endonuclease during the initial T-shift period and this increase continued up to 5 hours. After an optimum production level at 5 hours of T-shift, the enzyme yield also decreased. There may be at least two reasons for the reduction in the product yield at longer periods of T-shift: recombinant cell death due to excess product levels or high copy number, and a dilution of the culture by host cells. Consequently, T-shift

periods longer than 5 hours is harmful to the recombinant cell population and reduces the protein productivity.

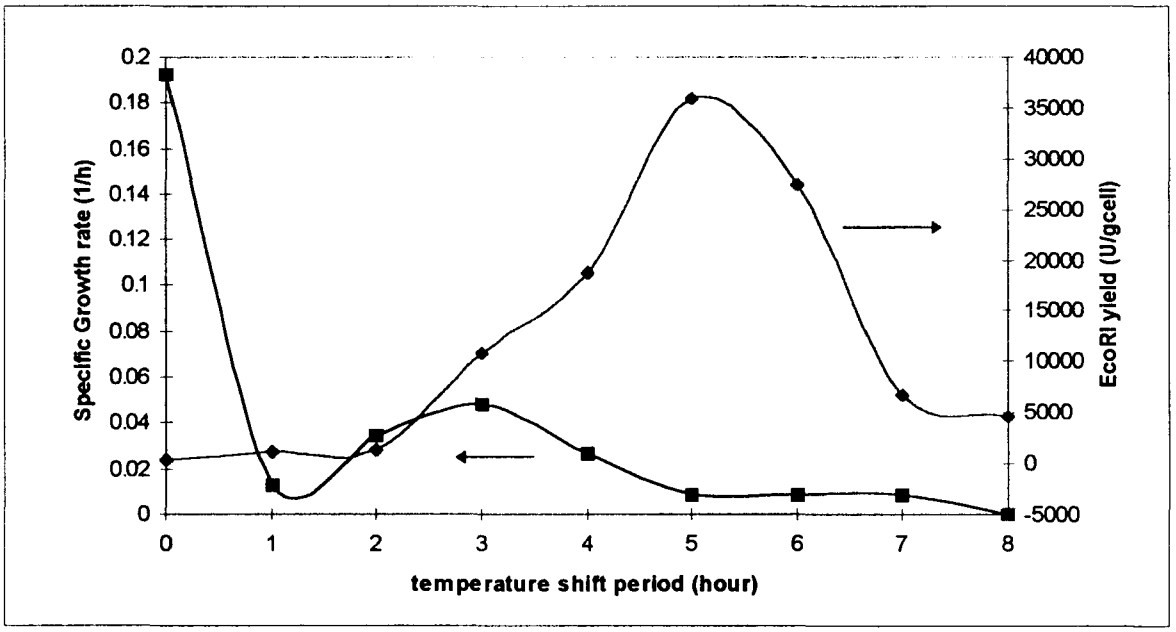


FIGURE 4.26. The change in the specific growth rates (μ) and *Eco*RI endonuclease production during T-shift period of 8 hours

4.4.1.2 The Effect of Cycling Time on Product Yield. The results reported in this section were obtained from experiments where *E.coli* M5248 (pSCC2) cells were subjected to elevated temperature (42°C) for different T-shift periods followed by a lowering of the temperature back to 32 °C. The effect of these varying time intervals on the optimal production of *Eco*RI endonuclease was examined to determine the cycling time for maximizing plasmid-coded product yield.

E.coli M5248 containing plasmid pSCC2 cells were grown at 30-32°C in supplemented L-broth to an absorbance value of 1.0 at 590 nm. The culture temperature was then raised to 42°C for specified periods of time between 1-5 hours; the temperature was then lowered back to 32°C, cell mass and enzyme production were determined. The time periods were limited to a maximum of five hours, since it was previously found that

the plasmid containing cell percentage dropped below 50 % after a five hour T-shift period. The optical densities were monitored and plotted against time (Figure 4.27). A second exponential phase was observed in the growth curves of the recombinant cultures showing that cells left the stationary phase they have entered because of an increase in the temperature, and continued to grow when 32°C was resumed.

The maximum specific growth rates calculated from the first and the second exponential phases of the recombinant cultures are given in Table 4.6. After lowering the temperature, cells continue to grow at one third of their original growth rates. The recombinant cells started growing at a slightly higher rate only after the 1- hour T-shift period. The growth rates calculated from the higher T-shift periods between 2-5 hours, were found to be similar to each other.

The *Eco*RI enzyme activity was determined at the end of each set of experiments in order to follow the effect of lowering the temperature back to 32°C. The protein amounts obtained after the first two runs corresponding to 1- and 2-hour T-shift periods followed by fermentation at 32°C were higher than the protein amounts obtained without lowering the temperature back to 32°C (Table 4.7). This shows that the cells immediately adapt themselves and continue to synthesize proteins. However , after the runs corresponding to 3-5 hour T-shift periods, the protein amounts remain constant. It is interesting to note that the activity of the *Eco*RI endonuclease is affected by the temperature-downshift. A decrease was observed in the total activity values eventhough the total protein amount stayed almost constant.

TABLE 4.6. The maximum specific growth rates (μ_m in hr^{-1}, T-shifts from 1-5 hrs)

T-shift (hr)	1	2	3	4	5
μ_m- 1st	0.329	0.346	0.330	0.303	0.328
μ_m - 2nd	0.097	0.086	0.085	0.080	0.086

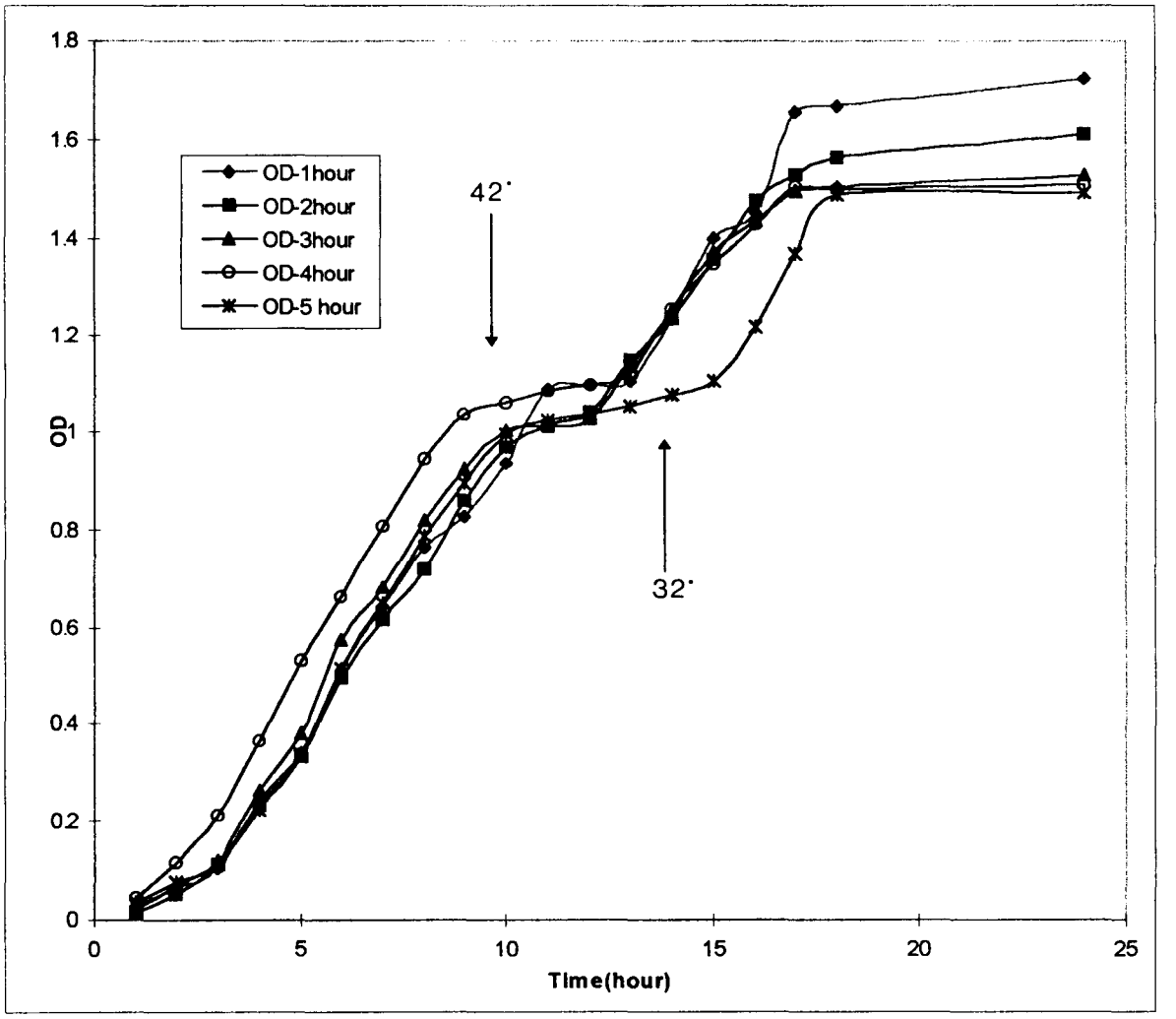


FIGURE 4.27. Growth behaviour of the recombinant cultures which were shifted from 32°C to 42°C between 1-5 hours, then shifted back to 32°C.

TABLE 4.7. The effect of temperature downshift to 32°C on the *Eco*RI endonuclease activity

T-shift period	Total Protein (mg)- 42°C	Total Protein (mg) - back to 32°C	Total Activity (U) - 42°C	Total Activity (U) -back to 32°C
1	118	214	$1.2x10^4$	$2.5x10^4$
2	151	235	$1.4x10^4$	$2.5x10^4$
3	328	354	$1.5 x10^5$	$1.2 x10^5$
4	492	484	$2.75 x10^5$	$2 x10^5$
5	524	516	$5.2 x10^5$	$4 x10^5$

diPasquantonio et al (1987) have investigated the effect of temperature-shifting on β-galactosidase production by the temperature-sensitive plasmid, pOU140, in *E.coli* CSH50 cells. They have concluded that a sustained thermal induction results in recombinant cell death and instability, while exposure to a run-away temperature for minimal time periods does not give sufficiently high product yields. However, the recombinant cells remain stable at intermediate cycling times, and the plasmid replication region is activated resulting in higher product yields. They have proposed that the specific activity of the enzyme could be improved if the cell culture temperature is raised from 31°C to 42 °C and then lowered back to 37°C, instead of a single temperature-upshift to 42°C. Promotion at the origin of replication leading to an increased copy number results in higher levels of β-galactosidase. A similar result may not be expected in the case of pSCC2 in *E.coli* M5248 since a maximization of the recombinant protein production is being sought. Table 4.7 shows that although higher *Eco*RI endonuclease yields were obtained by T-downshift only after the short (1-2 hour) T-upshift periods, these yields are still lower than those obtained by using longer single T-upshift periods.

4.4.2 Stability of Plasmid pPG430 in *E.coli 294* Cells

In the present section, the main aim was to determine the effect of various culture conditions on the recombinant plasmid stability and consequently on the *Eco*RI endonuclease productivity. In order to examine the effect of the presence of the recombinant vector on the growth rate of the host cells, the growth characteristics of the recombinant *E.coli* 294 cells were compared with those of native host cells in both minimal and complex media. Although minimal medium is known to reduce the growth rate and increase plasmid instability, it is difficult to follow the substrate consumption in most complex media (Kumar et al., 1991; Chou et al., 1994).

In the first part of this study, three parallel shake flask fermentations were performed, in two of which the recombinant cells were grown at 37°C in L-broth (complex

medium) under selective (with antibiotic) and non-selective (without antibiotic) environment to investigate whether the addition of antibiotics to the medium restricts growth, and increases the percentage of the plasmid containing cells. Cells were induced for the synthesis of *Eco*RI endonuclease at an absorbance value of 1.2 at 595nm by the addition of 0.1 mM IPTG. In the third shake flask, host cells were grown at 37°C also in L-broth. Optical densities, the number of viable cells and *Eco*RI activities were measured during cell growth. The growth curves are presented in Figure 4.28 and the growth rates given in Table 4.8 were calculated for host and recombinant cells. The host and the recombinant cells grown in non-selective environment completed the exponential phase faster than the recombinant culture grown in selective medium. The growth rates of host cells were higher than those of recombinant cells in all media tested. The maximum specific growth rate of recombinant cells cultured in non-selective medium was found to be sligthly higher than that in selective LB medium. This is in agreement with other findings in the literature (Ryu and Lee, 1987).

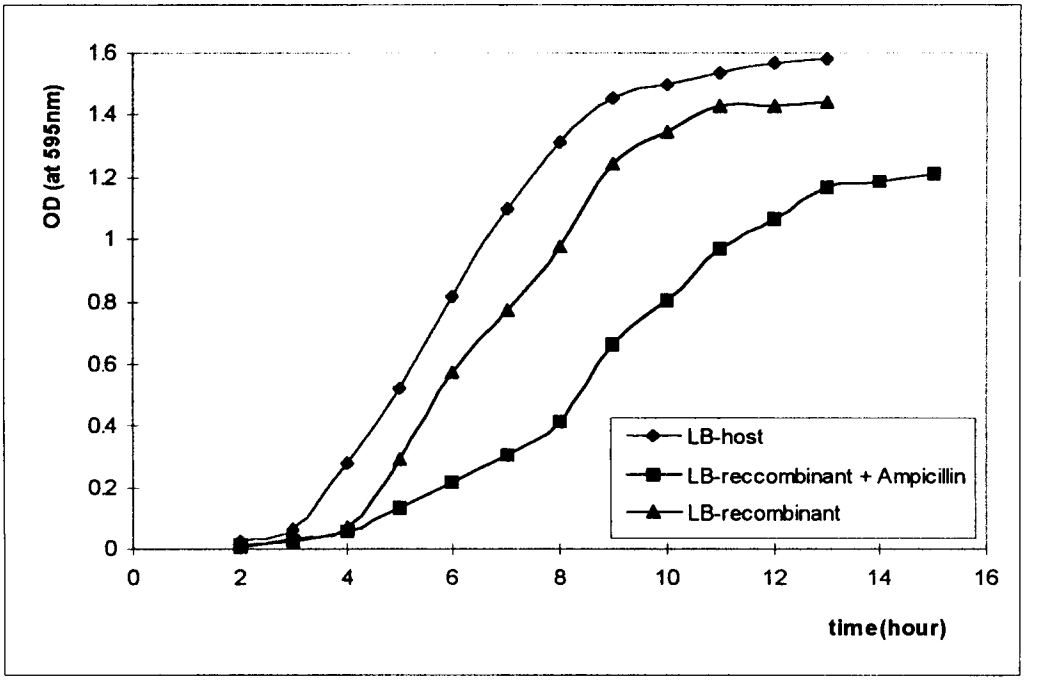


FIGURE 4.28. The growth curves of host and recombinant *E.coli* 294 cells in selective and non-selective complex media.

In the second part of the study, five parallel shake flask fermentations were performed, in three of which the recombinant cells were grown at 37°C in *minimal medium*. In the first shake flask fermentation, minimal medium was used without addition of aminoacids and ampicillin. In the second one, the medium was supplemented with aminoacids while in the third one, the medium was supplemented by both ampicillin and aminoacids. In the fourth shake flask fermentation, the host cells were grown at 37°C in minimal medium without addition of aminoacids whereas in the fifth one, the host cells were grown in minimal medium supplemented with aminoacids. The experiments were conducted as in the case of the ones completed in L-broth. The growth curves are presented in Figure 4.29 for five parallel shake flask fermentations. The addition of the aminoacids was found to be necessary to have meaningful growth of both the recombinant and the host cells. The addition of ampicillin was also found to have a slight restricting effect on the growth in minimal medium.

The maximum specific growth rates obtained both for complex and minimal medium fermentations are summarized in Table 4.8. The addition of the aminoacids to the minimal medium caused a significant increase in the maximum specific growth rates of the host and the recombinant cells, the growth rates were almost tripled in both cases. The growth rates of host cells were found to be higher than those of recombinant cells cultured in both media, probably reflecting the metabolic burden of the plasmid on the host cell. The growth rates were found to be higher in the rich medium than in the minimal medium.

TABLE 4.8. The Maximum Specific Growth Rates of Recombinant and Host Cell Cultures in Different Medium Composition (μ_m^- = host cells, μ_m^+ = recombinant cells, $\pm$ 3%)

	M9 Medium	**M9 Medium + Aminoacid + Ampicillin**	**M9 Medium + Aminoacid**	**LB Medium + Ampicillin**	**LB Medium**
μ^-_m	0.164	-------	0.429	-------	0.672
μ^+_m	0.098	0.307	0.318	0.520	0.549

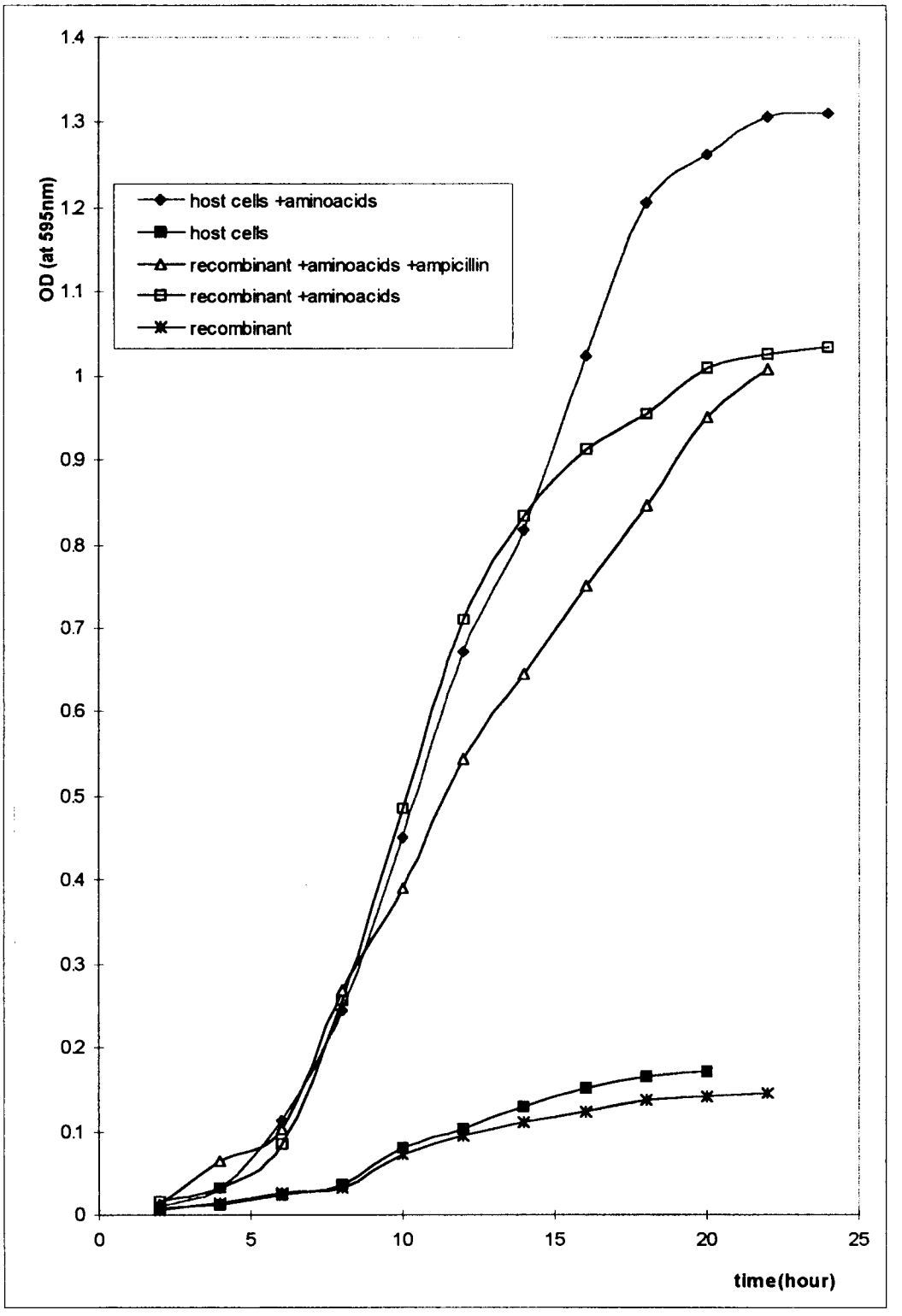


FIGURE 4.29. The growth curves of host and recombinant *E.coli* 294 cultures grown in M9 medium

The percentage of the plasmid containing cells were obtained from the ratio of the number of recombinant viable cells to the number of total viable cells and determined from cultures which were grown in non-selective medium. The changes in the percentage of the plasmid-containing cells in the recombinant cultures grown in complex and minimal media are shown in Figure 4.30. The decrease in the stability of the plasmid is higher in minimal medium than in L-broth. The percentages of the cells containing plasmid determined in the minimal and complex media supplemented with ampicillin were higher than the non-selective cultures. In L-broth, the addition of antibiotics could increase the plasmid-containg cell percentage by only 5%, therefore the plasmid was found to be quite stable in non-selective complex medium (Table 4.9). Antibiotics addition to the medium can be a problem in large scale operations, since a large amount of antibiotics will be required. Therefore *E.coli* 294 (pPG430) cells can also be used for large scale fermentations without addition of ampicillin. The presence of antibiotics was found to be necessary to decrease the concentration of the mixed cell populations in the use of the minimal medium, since the percentage of the plasmid containing cells were increased by more than 15% by the addition of ampicillin to the environment. Even in the presence of ampicillin in the environment, the plasmid bearing cell percentage does not exceed 74% at the late exponential phase in minimal medium. This was 89% in L-broth. Therefore, the recombinant cells when cultured in complex medium were found to be more stable than in minimal medium.

TABLE 4.9. The percentage of plasmid containing cells at the late exponential phase in different selective and non-selective media

	M9 Medium + Ampicillin	**M9 Medium**	**LB Medium + Ampicillin**	**LB Medium**
%Plasmid Bearing Cells	74.8	59.4	89.1	83.7

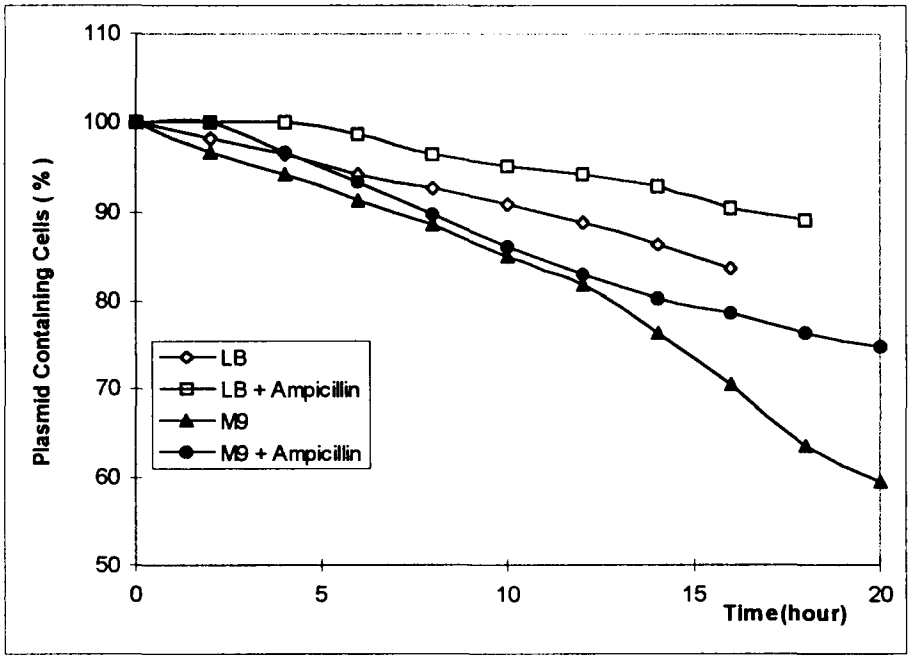


FIGURE 4.30. Effect of medium composition on the percentage of plasmid bearing cells

In Table 4.10, the growth rate differences are calculated for the recombinant and the host cells both in minimal and LB medium. There is a large and quantitatively similar difference in the growth rates of the host and recombinant cells. Possible conclusion which can be drawn from this large difference in the growth rates, is that the instability of the plasmid was due to the growth rate differences rather than segregational instability (Cooper et al., 1985; Davidson et al., 1990).

TABLE 4.10. Comparison of the Growth Ratios and the Growth Rate Differences of Recombinant (μ^{+}_{max}) and Host (μ^{-}_{max}) Cells

Medium	μ^{-}_{max}	μ^{+}_{max}	$\mu^{-}_{max} - \mu^{+}_{max}$	$\mu^{+}_{max} / \mu^{-}_{max}$
Minimal	0.429	0.307	0.122	0.716
LB	0.672	0.520	0.152	0.774

The maximum *Eco*RI enzyme activity was found when the cells were cultured in L-broth with the addition of antibiotics. Minimal medium was found to result in a decrease in the enzyme productivity probably due to the lower growth rate of the cells and the problem of plasmid instability (Table 4.11).

TABLE 4.11. The effect of different culture media on the *Eco*RI enzyme acivity

*Eco*RI	M9 Medium + Ampicillin	M9 Medium	LB Medium + Ampicillin	LB Medium
Total Activity (U)	6.0×10^6	$5.25 \times \times 10^6$	1.12×10^7	9.7×10^6
Specific Activity (U/mg)	15619	12987	26168	23317

4.4.3 Investigation of the Stabilities of Plasmids pPG430 and pSCC2 After Succesive Generations

The plasmid stabilities of both recombinant *E.coli* cells were also investigated after successive cultures.

In each set of experiments, two parallel shake flask fermentations were performed. Recombinant *E.coli* 294 and *E.coli* M5248 cells were inoculated separately in 100 ml of LB medium in the shake flasks. Recombinant cell cultures were grown at 37°C and 30°C respectively in orbital shakers until they reached their late exponential phase, then 1ml of sample was taken from each culture and transferred into two shake flasks containing 100 ml of LB medium. The cultures were again grown until the late exponential phase was reached.

This process was repeated until 20 batch cycles were completed. The total viable cells and the number of recombinant viable cells were determined at the end of the each batch fermentation (Section 3.3.7). Figure 4.31 shows the variation in the percentage of plasmid containing cells during 20 successive batches.

E.coli 294 cells containing plasmid pPG430 were found to be more stable in maintaining their plasmids. Even after the 18th batch, the plasmid percentage in the culture was higher than 50%, whereas in the case of *E.coli M5248* cells bearing plasmid pSCC2, the percentage of the recombinant cells was found to be smaller than 50% after the 5th batch. It is therefore advisable to use the freshly transformed cells for *E.coli M5248* (pSCC2) fermentations. A similar study was undertaken by Nagahari et al (Imanaka et al., 1981) who have investigated the change in the plasmid stability on a recombinant plasmid pBR322-T.leu in *E.coli* C600 cells. After 25 batches, they have found that the plasmid stability decreases from 97 % to 45 %.

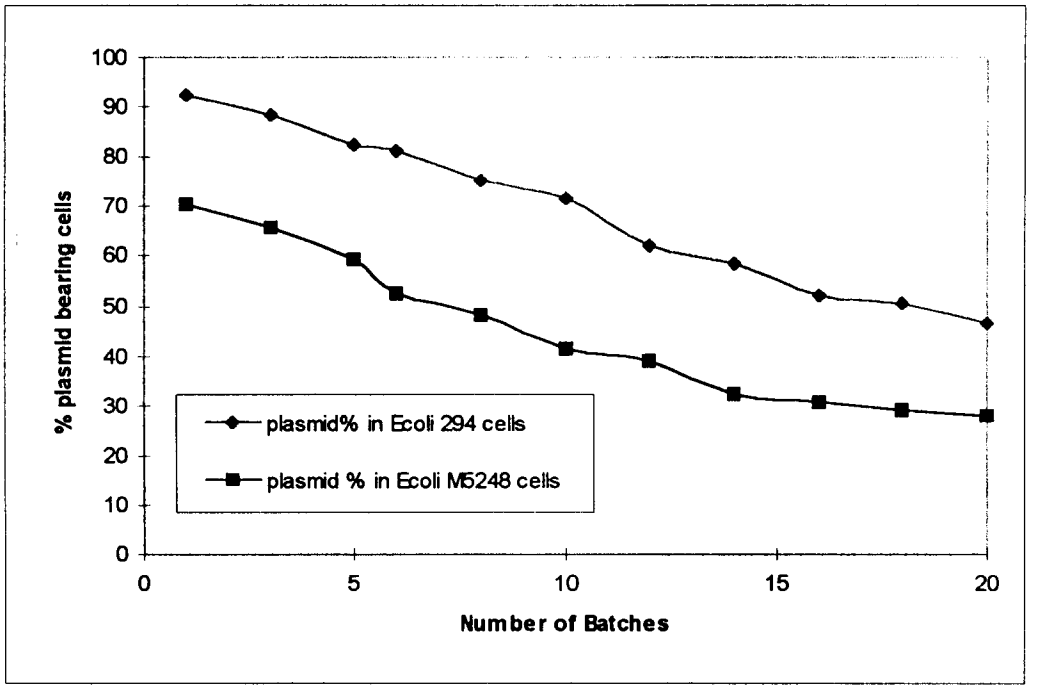


FIGURE 4.31. The variation in the stabilities of plasmid pSCC2 in *E.coli* M5248 and pPG430 in *E.coli* 294 during 20 successive batches

4.4.4 Mathematical Methods in the Analysis of Plasmid Stability

The specific growth rates of plasmid-containg cells and plasmid-free cells were compared for both *E.coli* 294 and M5248 and the difference in the growth rates were calculated by using the model presented by Cooper et al. (1987). The model is based on the assumption that the overall causes of plasmid instability can be described either by the segregational instability of the plasmid or the growth rate differences between the plasmid-free and the plasmid-bearing cells, or by both. They have developed a technique for calculating the proportion of plasmid free cells assuming there are three different patterns that unstable microbial populations can follow. These include the cases where the growth rate difference is much greater , smaller or equal to the segregational instability.

The mathematical model they have used for continuous fermentation systems is applied to batch fermentation in the present work. If X_+ and X_- represent the concentrations of plasmid-bearing and plasmid-free cells respectively, then the mathematical model describing the change in the concentration of each population with time is as follows:

$$\frac{dX_+}{dt} = \mu_+ \cdot X_+ - R \cdot X_+ \qquad (4.1)$$

$$\frac{dX_-}{dt} = \mu_- \cdot X_- + R \cdot X_+ \qquad (4.2)$$

where μ_+ and μ_- are the growth rates of X_+ and X_- , and R is the segregation rate. The actual values of X may be expressed in terms of the fractions of the total cell population;

$$p_+ = \frac{X_+}{X_+ + X_-} \quad \text{and} \quad p_- = \frac{X_-}{X_+ + X_-} \qquad (4.3)$$

Equations (4.1) and (4.2) can also be written in terms of the mass fractions p_+ and p_- :

$$\frac{dp_+}{dt} = \mu_+ \cdot p_+ - R \cdot p_+ \qquad (4.4)$$

$$\frac{dp_-}{dt} = \mu_- \cdot p_- - R \cdot p_- \qquad (4.5)$$

Adding the two equations gives

$\mu_+ p_+ + \mu_- p_- = 0$

If the total cell concentration ($X_+ + X_-$) is assumed constant at the early stationary phase, and using the definitions

$d\mu = \mu_- - \mu_+$ (4.6)

$p_+ + p_- = 1$ (4.7)

it can be shown that

$\mu_+ = - p_- d\mu$ (4.8)

and

$\mu_- = - p_+ d\mu$ (4.9)

Substituting Equations (4.8) and (4.9) into equation (4.5) gives

$$\frac{dp_-}{dt} = \mu_- \cdot p_- + R \cdot p_+ \quad (4.5\ a)$$

Considering the fact that $p_+ + p_- = 1$ and $\frac{dp_-}{dt} = -\frac{dp_+}{dt}$

and substituting into equation (4.5.a)

$$\frac{dp_+}{dt} = -\mu_-(1-p_+) - R \cdot p_+ \qquad (4.5\ b)$$

Substituting from Equations (4.8) and (4.9) and rearranging gives

$$\frac{dp_+}{dt} = p^2{}_+ \cdot d\mu - p_+ \cdot (d\mu + R) \qquad (4.10)$$

Assuming that R and dμ are constant, this differential equation is of the Bernouille form and can be solved by letting $(1 / p_+) = z$ (Dwight, 1961). The solution has been rearranged to yield:

$$p_- = \frac{(p_{-0}\, d\mu + R)\, e^{(d\mu + R)\, t} - R\,(1 - p_{-0})}{(p_{-0}\, d\mu + R)\, e^{(d\mu + R)\, t} + d\mu\,(1 - p_{-0})} \qquad (4.11)$$

where p_{-0} is the fraction of plasmid-free cells at t= 0.

Equation (4.11) can be simplified by considering the fact that $p_{-0} << 1$:

$$p_- = \frac{(p_{-0}\, d\mu + R)\, e^{(d\mu + R)\, t} - R}{(p_{-0}\, d\mu + R)\, e^{(d\mu + R)\, t} + d\mu} \qquad (4.11\ a)$$

The three cases considered by Cooper et al (1987) can be expressed mathematically as follows:

1. $d\mu >> R$
2. $d\mu \leq R$
3. $d\mu < 0$ and $|d\mu| >> R$

The third case need not be considered in the present work, since this is the case where the recombinant cells have a growth rate advantage to the host.

Case I : $d\mu >> R$

In the first case, the R term in the exponent is negligible and the denominator is approximated by $d\mu$. Then equation (4.11a) becomes:

$$p_- = (p_{-0} + R / d\mu) e^{+d\mu t} - R / d\mu \qquad (4.12)$$

When the logarithm of the fraction of plasmid-free cells is plotted against time, in the linear region of the plot $t > 1/ d\mu$ and Equation (4.12) reduces to

$$p_- = (p_{-0} + R / d\mu) e^{+d\mu t} \qquad (4.12a)$$

and the slope and the intercept can be expressed as

slope = $d\mu$ intercept= $\ln (p_{-0} + R / d\mu)$

Case II: $d\mu < R$

In the second case, the p_{-0} $d\mu$ term can be neglected and a binomial expansion yields: Equation becomes:

$$p+ = \left(1+\frac{d\mu}{R}\right)\cdot e^{-(d\mu+R)\cdot t} + \frac{d\mu}{R}\cdot\left(1+\frac{d\mu}{R}\right)\cdot e^{-2(d\mu+R)\cdot t} \qquad (4.13)$$

In this case, a plot of the logarithm of plasmid-bearing cells against time is obtained, and in the exponential region where $t > 1/ (d\mu + R)$, Equation (4.13) becomes:

$$p_+ = (1 + d\mu / R) e^{-(d\mu + R) t} \qquad (4.13a)$$

with

Slope = $-(d\mu + R)$ Intercept = $\ln (1 + d\mu / R)$

The plots and the equations obtained from linear regression analysis of data for Case I and II for both *E.coli* strain are shown in Figures 4.32-37. The results obtained for each of the recombinant cultures are given in Table 4.12. Since the temperature-shift was applied to the *E.coli M5248* culture at the late exponential phase, it has no effect on the growth during exponential phase. Out of eight set of experiments , two were analyzed. Plasmid stabilities of *E.coli M5248* (pSCC2) and *E.coli 294* (pPG430) were experimentally determined for both strains and the result indicated that the growth rate differences were the possible cause of this instability (Section 4.4.1- 4.4.3). In this section, the mathematical model presented as Case I and II were applied to both recombinant systems. The calculated differences for Case I were found to be close to the experimental values supporting the conclusion derived earlier. Experimental growth rate differences of *E.coli M5248* (pSCC2) cells were 0.272 h^{-1} for 3 and 5 hour periods of temperature-shift, while calculated differences were found to be 0.272 and 0.238 hr^{-1} for the same runs. For *E.coli 294* (pPG430) cells cultured in minimal and complex media, the experimental differences were 0.122 and 0.152 hr^{-1} , while the calculated ones were found as 0.157 and 0.149 hr^{-1} , respectively. The mathematical model used indicates that the plasmid instability is likely to be due to the growth rate differences for both recombinant systems. This is also seen in the relative values of $d\mu$ and R, especially in the case of *E.coli 294* (pPG430) in LB medium.

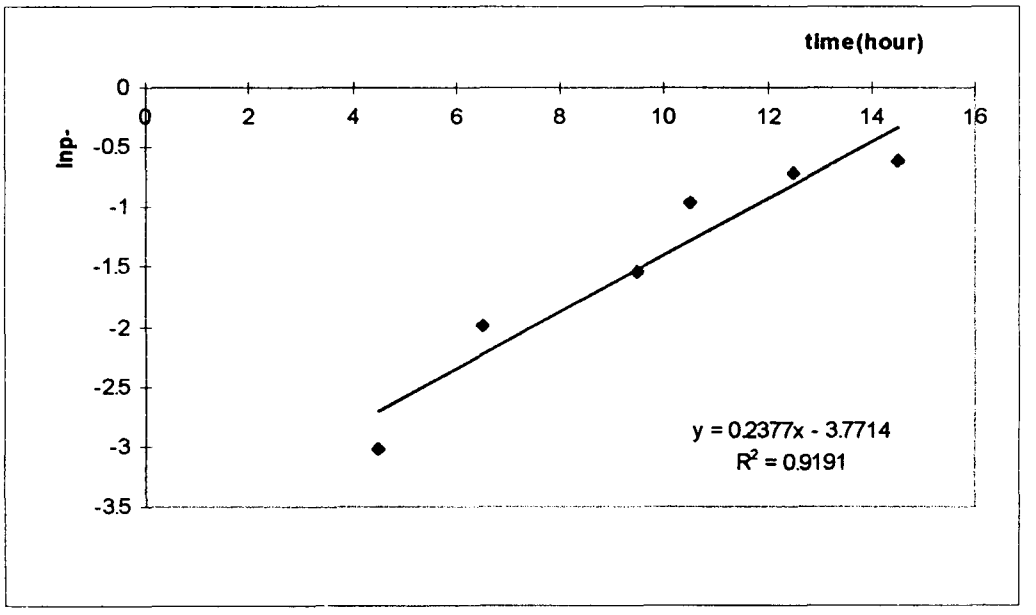


FIGURE 4.32. Case I: ln p_- versus time plot for *E.coli* M5248 exposed to 3 hours T-shift

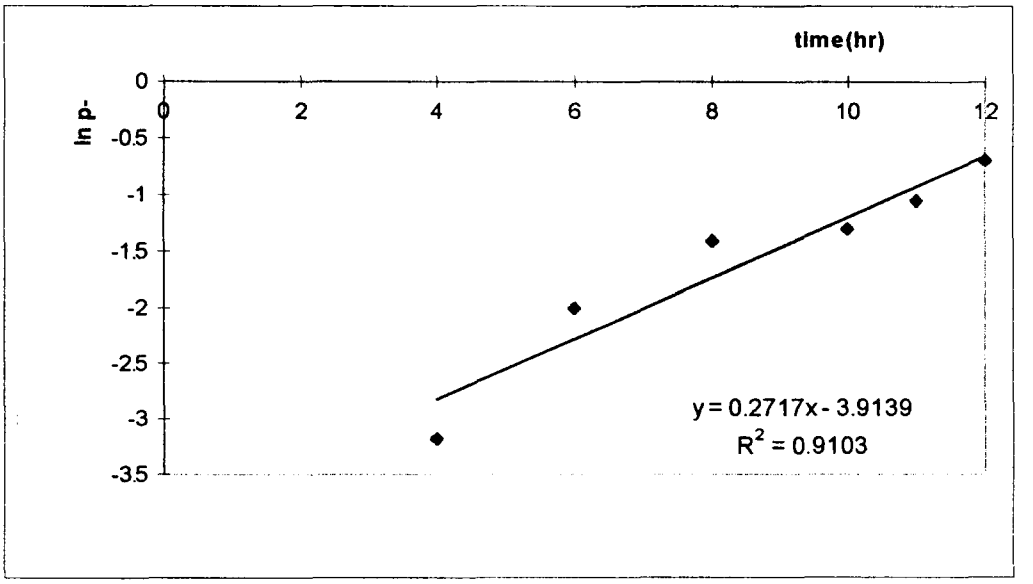


FIGURE 4.33. Case I: ln p_- versus time plot for *E.coli* M5248 exposed to 5 hours T-shift

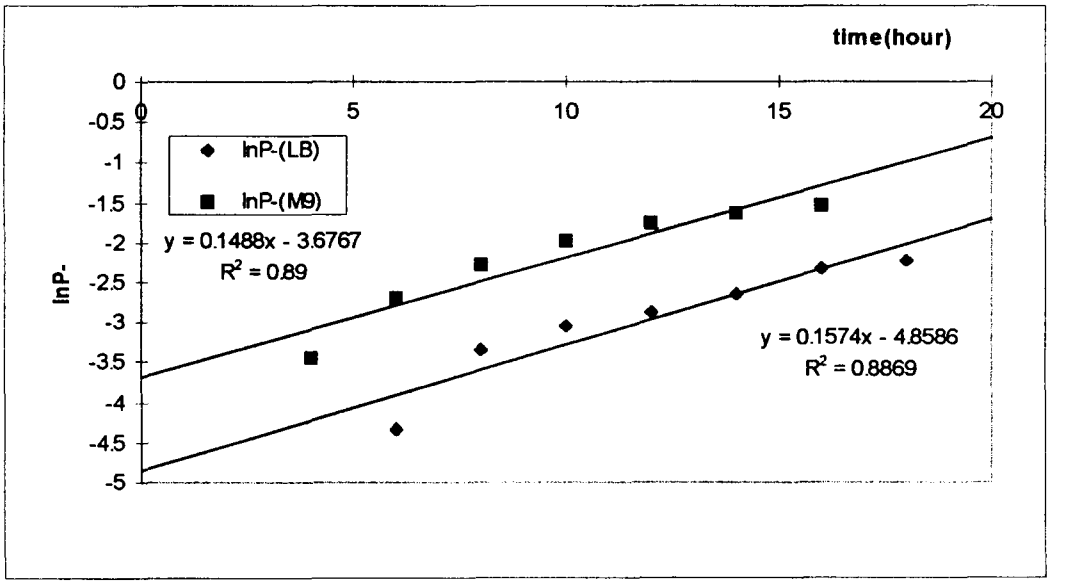


FIGURE 4.34. Case I: ln p_- versus time plot for *E.coli* 294 cultured in M9and LB media

The calculated differences that are given in Table 4.12 for Case II were found to be irrelavant to the experimantal values supporting the idea that $d\mu < R$ can not be the case for the two recombinant systems studied.

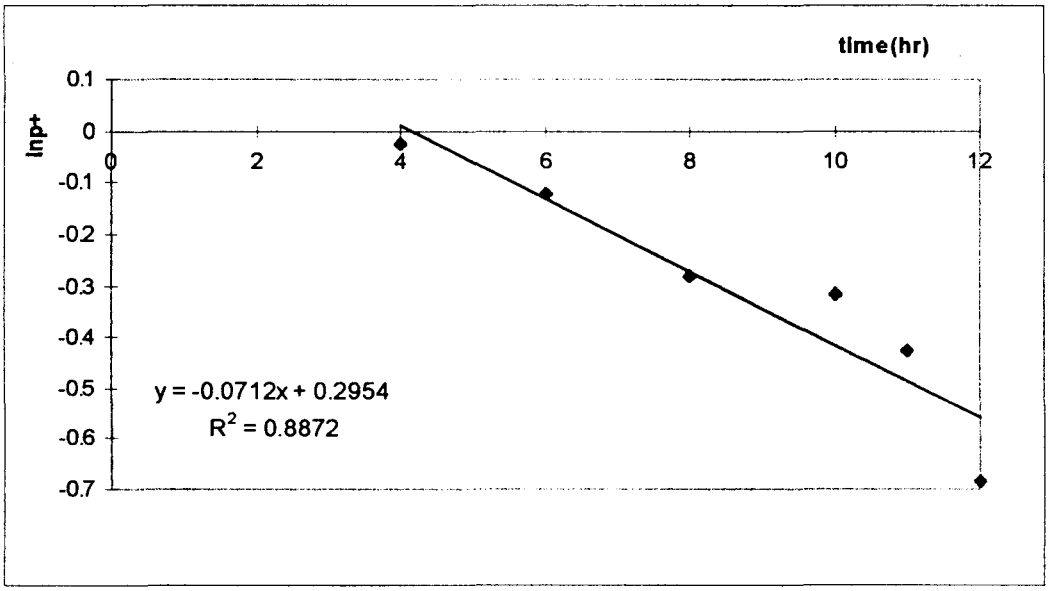


FIGURE 4.35. Case II: ln p_+ versus time plot for *E.coli* M5248 exposed to 3 hours T-shift

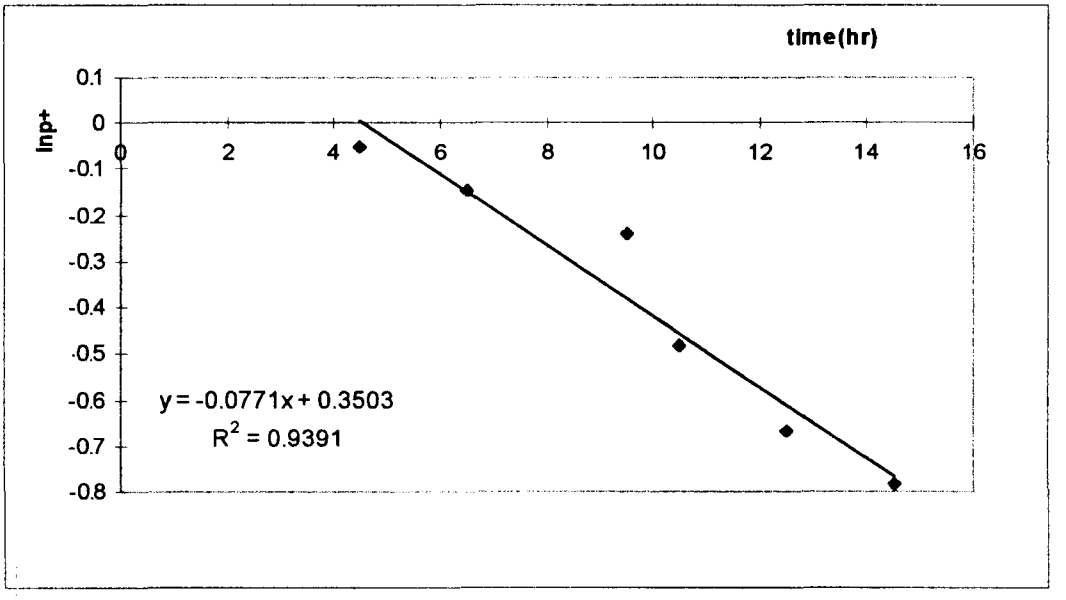


FIGURE 4.36. Case II: ln p_+ versus time plot for *E.coli* M5248 exposed to 5 hours T-shift

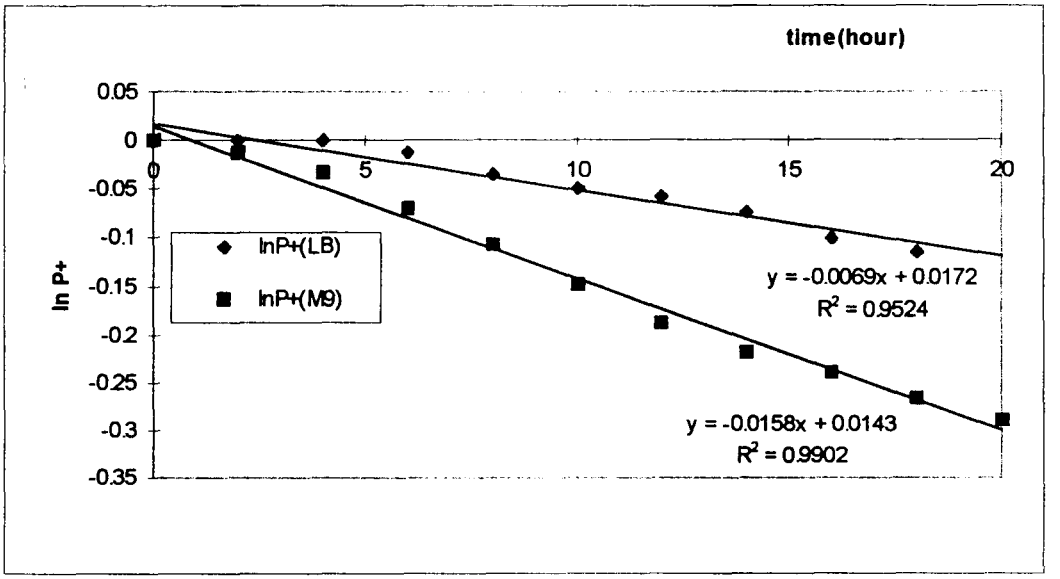


FIGURE 4.37. Case II: ln p_+ versus time plot for *E.coli* 294 cultured in M9and LB media

TABLE 4.12. Comparison of experimental and calculated values of the growth rate differences and segregational instability

	3-hours T-shift		**5-hour T-shift**	
***E.coli M5248* (pSCC2):**	calculated	experimental	calculated	experimental
dμ>>R **dμ**	0.272	0.272	0.238	0.272
R	0.00665		0.005472	
***E.coli M5248* (pSCC2):**	calculated	experimental	calculated	experimental
dμ<R **dμ**	0.002821	0.272	0.02278	0.272
R	0.05298		0.05431	
	LB Medium		**Minimal Medium**	
***E.coli 294* (pPG430):**	calculated	experimental	calculated	experimental
dμ>>R **dμ**	0.149	0.152	0.157	0.122
R	0.00377		0.00122	
***E.coli 294* (pPG430):**	calculated	experimental	calculated	experimental
dμ<R **dμ**	0.000118	0.152	0.00023	0.122
R	0.006782		0.01557	

In general, two major parameters have been studied by a number of investigators; the probability of plasmid loss **due to segregation** during cell division and **the difference in the growth rate** between the recombinant and the reverted. However little attempt has been made in these studies to quantify the magnitute of either the probability of plasmid loss or the difference in specific growth rate (Cooper et al., 1987; Davidson et al., 1989; Kumar et al, 1991).

The productivity of a bioreactor employing recombinant cells is largely affected by the degree to which the plasmid-free cells (p_-) are generated and propagated. This phenomena could be responsible for further complications in the scale up which is required for commercializing the recombinant products. The p_- cells are generated from plasmid-harboring p_+ cells by **segregational instability** which is caused by defective partitioning of the plasmids between the daughter cells during cell division (Kumar et al., 1991). The p_+ cells usually grow more slowly than the p_- cells because the p_+ cells have to synthesize more DNA, mRNA, and protein. This leads to a lower maximum growth rate, μ_{max} for p_+ cells. The **growth rate** of p_+ cells also depends upon the toxicity of the coded proteins and the strength of the promoters. Thus, once generated in the reactor, p_- cells may propagate rapidly, leading to a mixed population with the p_- proportion of population increasing with each generation, leading in turn to poor bioreactor process.

4.5 The Effect Of Substrate Concentration on the Growth and *Eco*RI Production of Recombinant *E.coli* 294 and *E.coli* M5248 Strains

Specific growth rate is the major parameter that determines the growth characteristics of cell cultures. Quantitative analysis of cell growth is based on the variation of biomass with time in the batch fermentation results presented in this part of the

study. Although a complex medium is to be preferred for favoring cell growth, it is difficult to examine the effect of substrate utilization by the organism during the growth cycle in a complex medium. Therefore minimal medium supplemented with aminoacids and glucose was used as the fermentation medium.

The effect of the change in substrate concentration on growth and product formation capability of the cells was investigated in minimal medium contaning five different initial glucose concentrations, namely 1, 5, 10, 15 and 30 g/L. Shake flasks containing 1L of minimal medium each with different glucose concentrations were inoculated with 25ml of a preculture at its late exponential growth phase. They were then placed in orbital shakers at 30°C and 37°C for *E.coli* M5248 and *E.coli* 294 strains respectively. At the late exponential phase of growth, cell cultures were induced for product formation under the specified conditions explained in the previous section (Section 4.1). Aliquots of the cell suspension were taken at the desired time intervals an and cell growth was followed by determining both the optical density and the dry weight of cell cultures. Residual glucose and *Eco*RI endonuclease activity in each aliquot were measured to follow substrate utilization and product formation.

The variations in the dry cell weight and glucose concentration of each cell culture are plotted versus time in Figures 4.38-4.47. A lag phase ranging between 3-6 hours is observed for each fermentation broth. During the lag phase, glucose consumption is almost constant, then a rapid decrease is observed as the cells enter their exponential growth phase. Substrate consumption in the media contaning 1 to 15 g/L glucose was complete, while around one quarter of the initial glucose was not consumed in the culture medium with 30 g/L of initial glucose concentration. Corresponding trends were observed in the time profiles of dry cell weights (Table 4.13). Dry cell weights determined at the stationary phase were found to be very low for both recombinant cell cultures with an initial glucose concentration of 1g/L. The cell mass increased by a factor of 1.8 in *E.coli* M5248 and by a factor of 2.2 in *E.coli* 294 when an initial glucose concentration of 5 g/L was used. The use of higher concentrations (10-15 g/L) did not result in any significant increase in the final cell mass, which reached 2.8 and 3.0 g/L at the end of the fermentation with initial glucose concentrations of 10 - 15 g/L respectively for *E.coli* M5248 cells. These values were 3.1 and 3.25 g/L for *E.coli* 294 cells for the same initial substrate concentrations. The final cell

mass obtained for *E.coli* M5248 with higher initial glucose concentration (30 g/L) remained the same, while a small increase was observed in the case of *E.coli* 294.

TABLE 4.13. Comparison of the fermentation results of both recombinant strains

Strain	S_0 (g/L)	S_{final} (g/L)	X_{final} (g/L)	μ_{max} (hr^{-1}) (± 3%)
E.coli 294	1	0.05	1.41	0.187
	5	0.11	3.05	0.289
	10	0.36	3.08	0.305
	15	0.69	3.25	0.316
	30	7.64	3.91	0.246
E.coli M5248	1	0.00	1.36	0.192
	5	0.05	2.55	0.233
	10	0.22	2.79	0.289
	15	0.47	3.03	0.300
	30	6.62	3.12	0.256

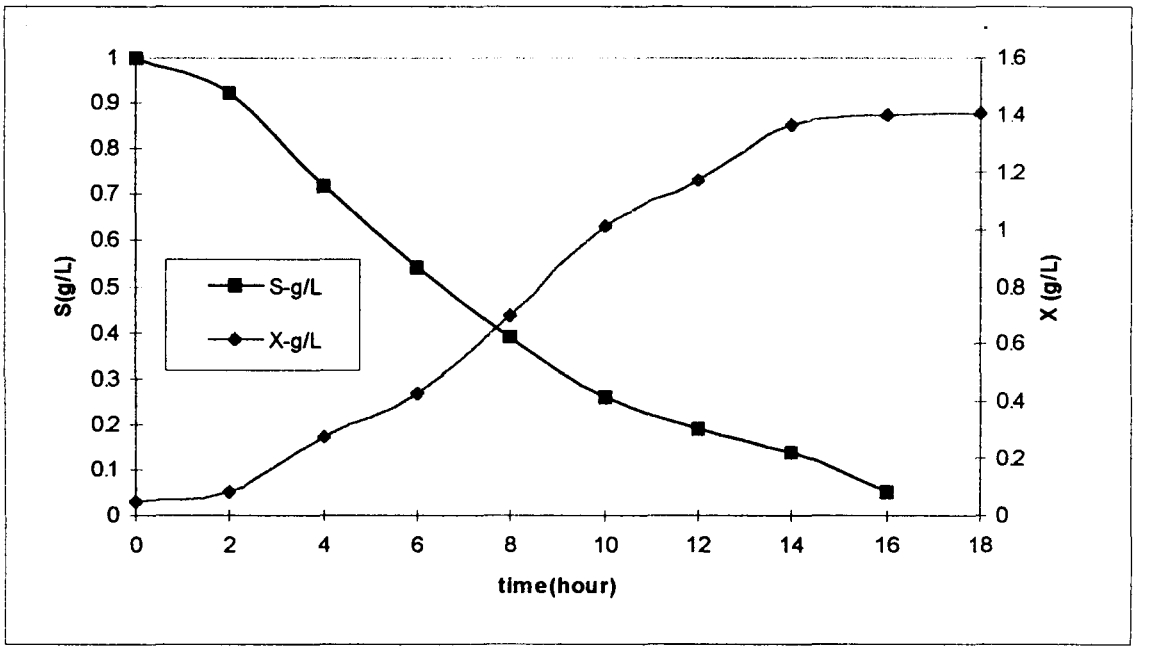


FIGURE 4.38. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 1 g/L)

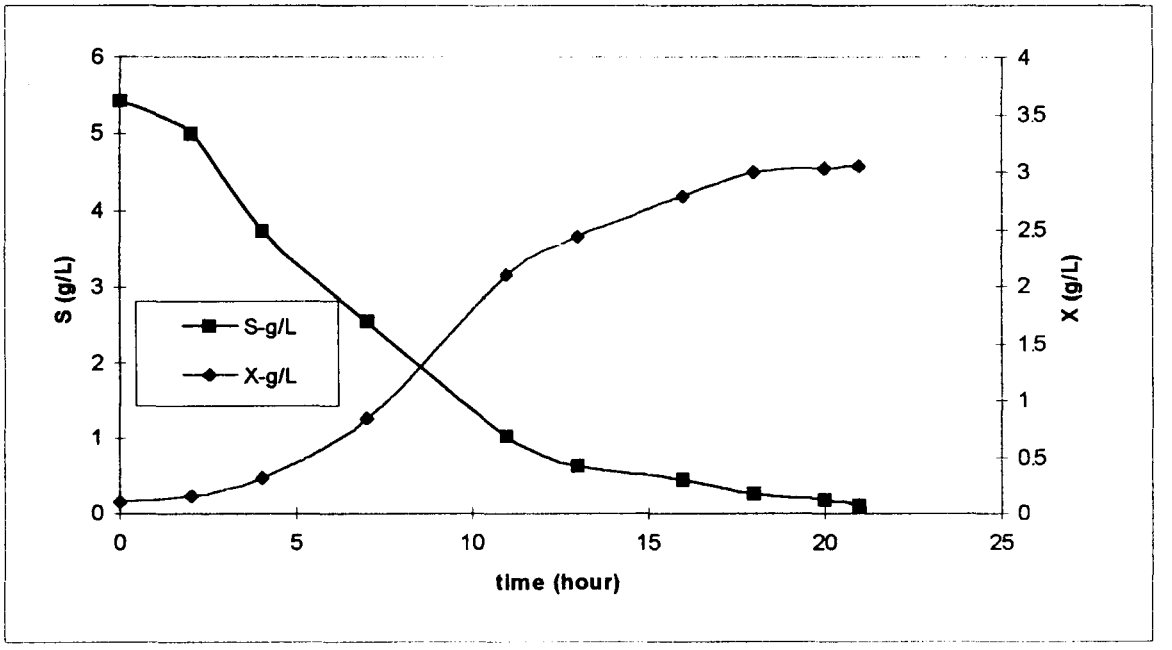


FIGURE 4.39. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 5 g/L)

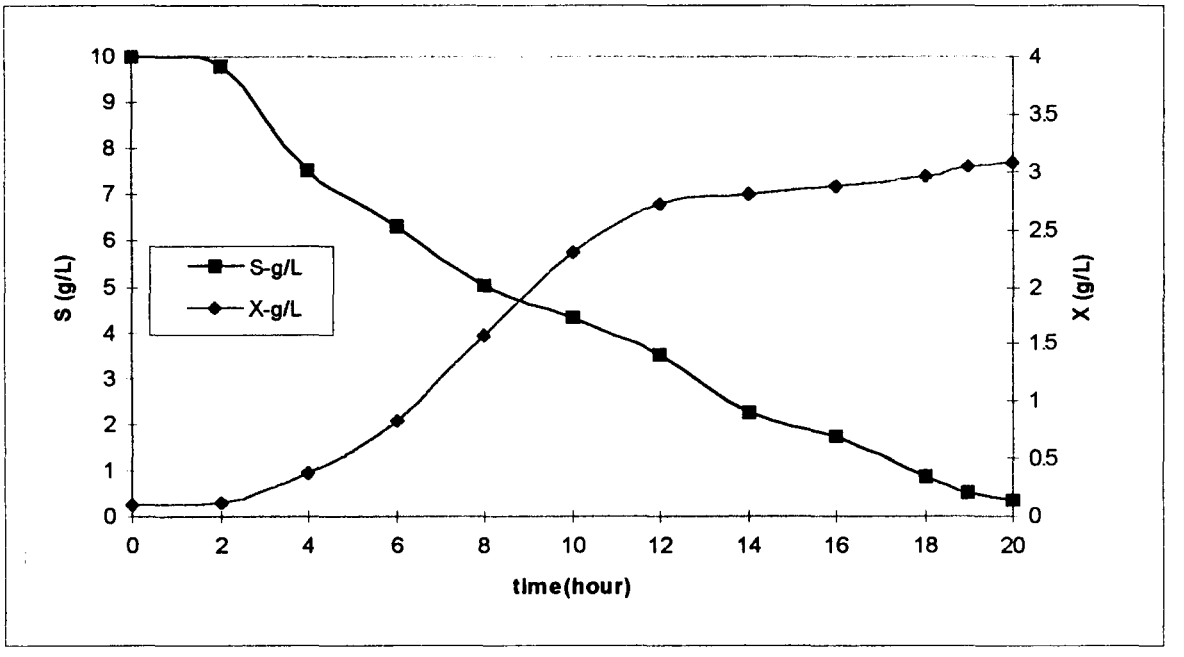


FIGURE 4.40. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 10 g/L)

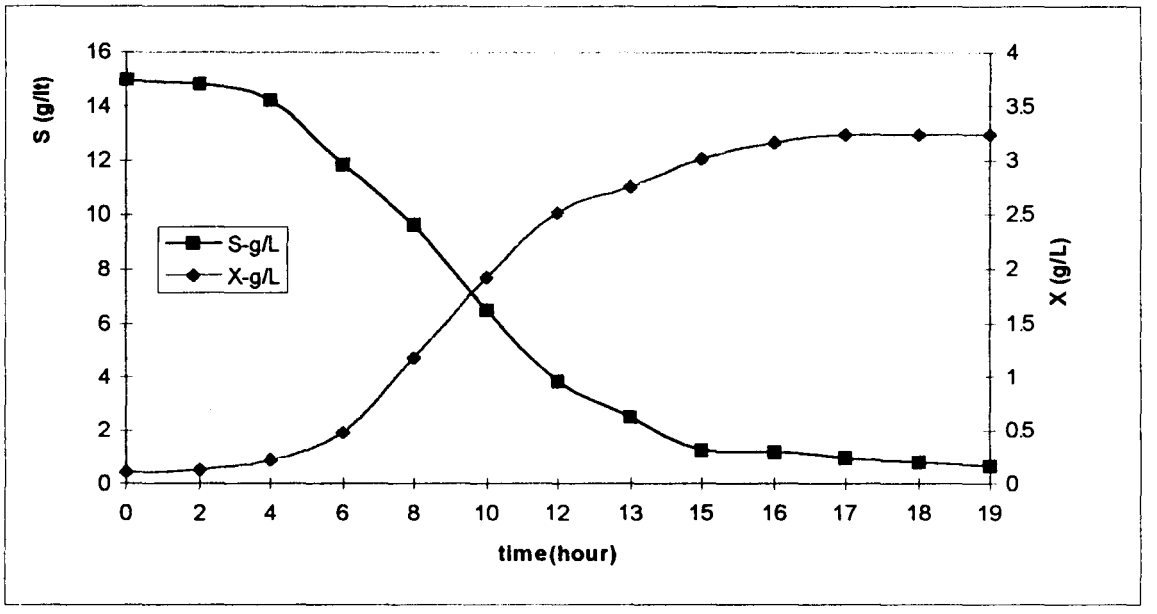


FIGURE 4.41. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 15 g/L)

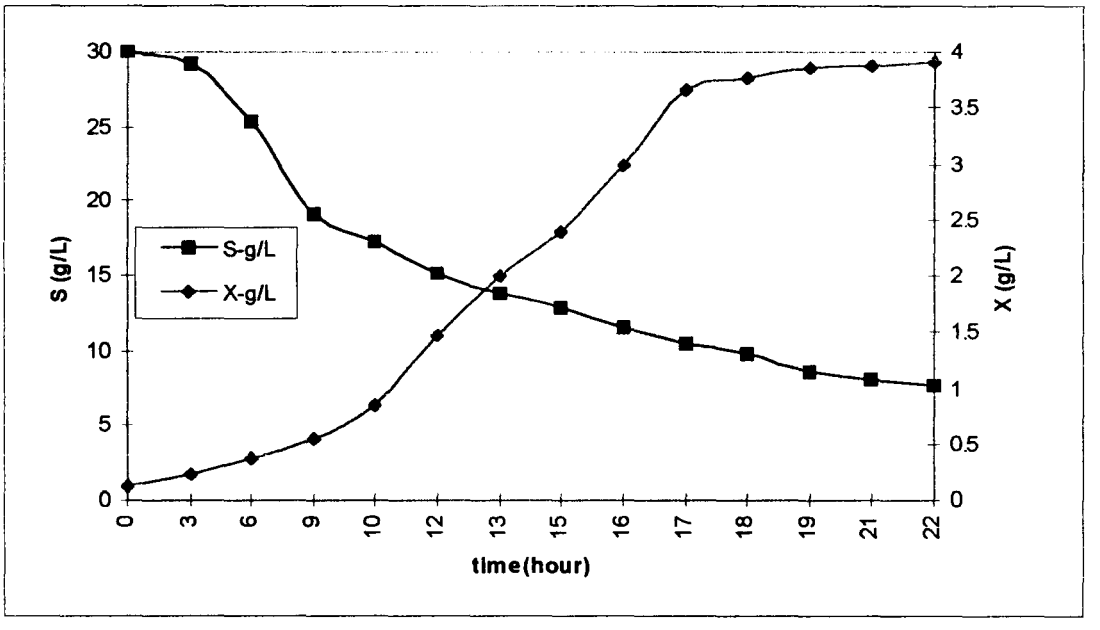


FIGURE 4.42. Time profiles of cell mass and residual glucose concentration for *E.coli* 294 cells in minimal medium (S_0 = 30 g/L)

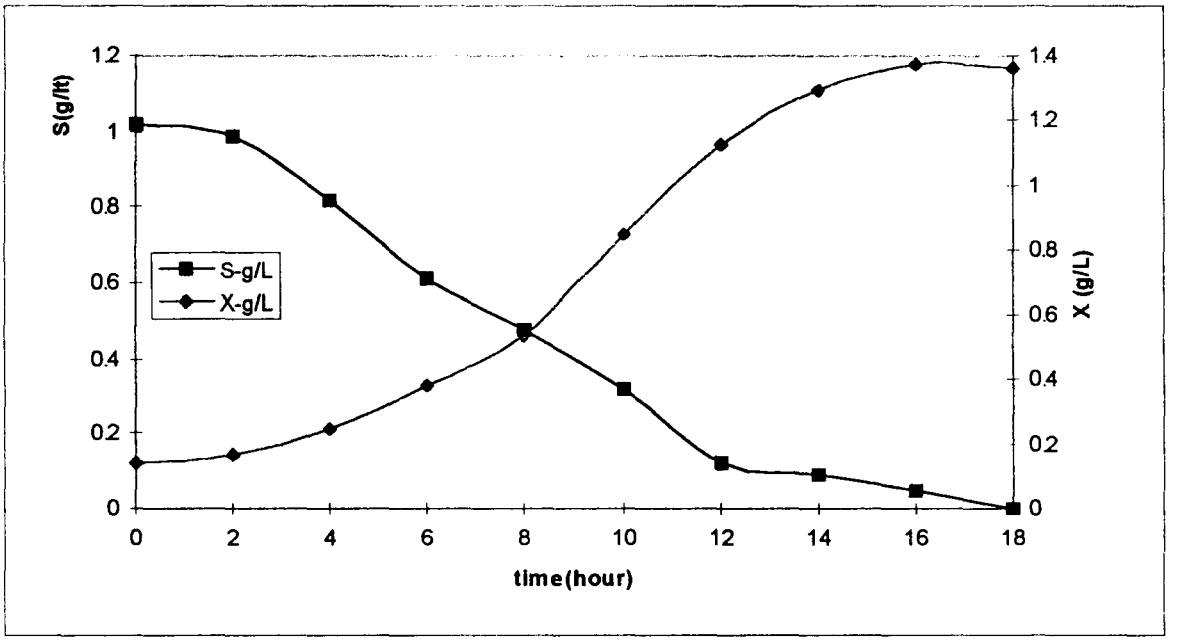


FIGURE 4.43. Time profiles of cell mass and residual glucose concentration for *E.coli*M5248 cells in minimal medium (S_0 = 1 g/L)

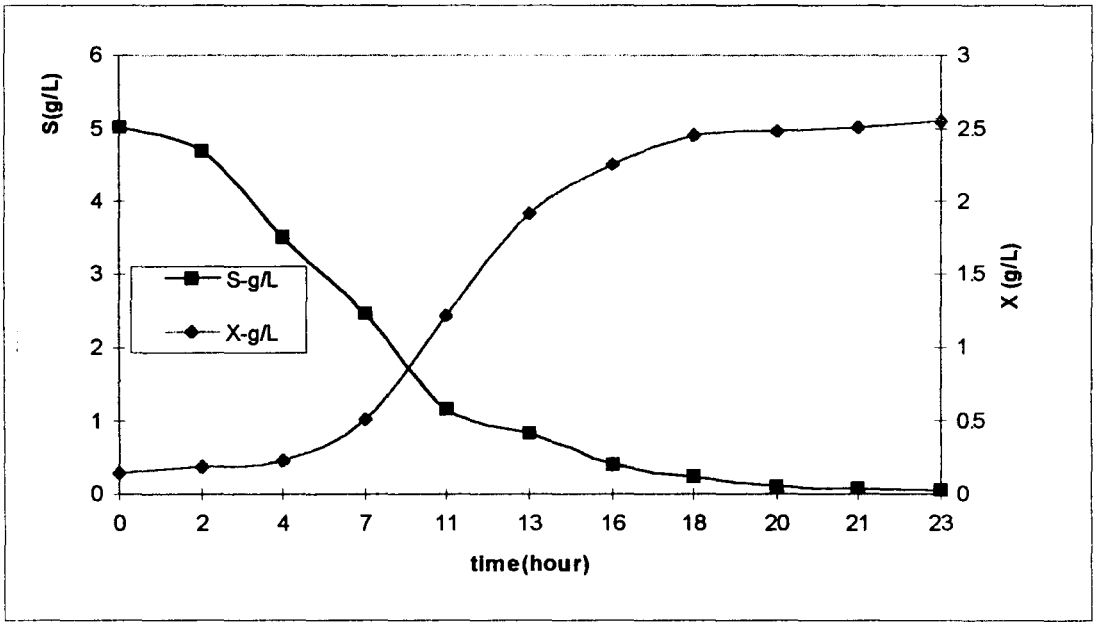
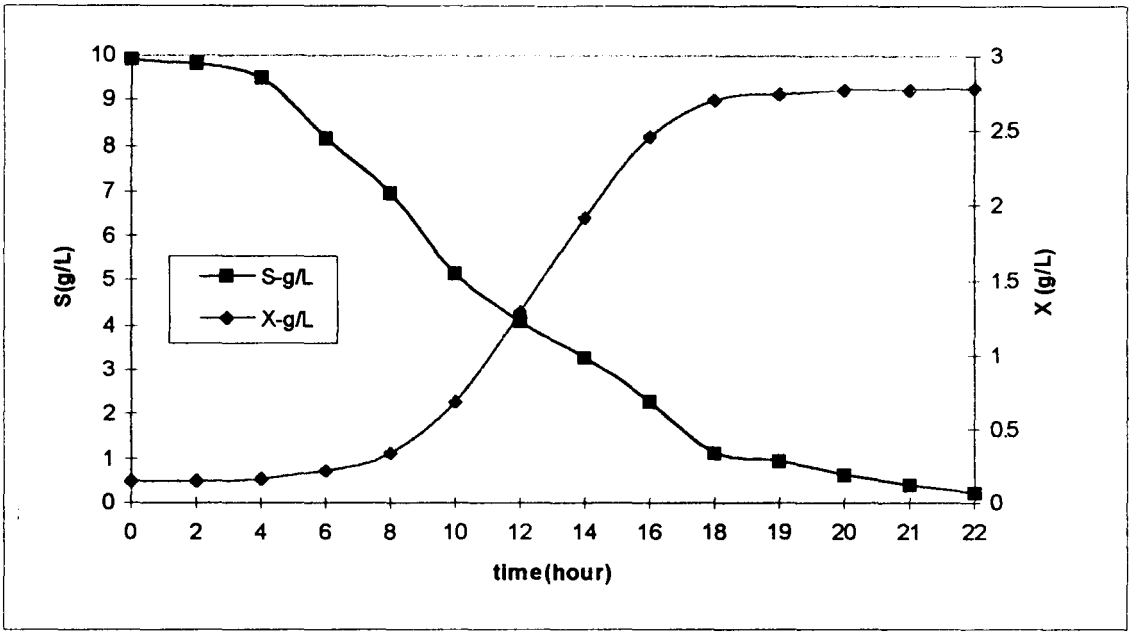


FIGURE 4.44. Time profiles of cell mass and residual glucose concentration for *E.coli*M5248 cells in minimal medium (S_0 = 5 g/L)



FiIGURE 4.45. Time profiles of cell mass and residual glucose concentration for *E.coli*M5248 cells in minimal medium (S_0 = 10 g/L)

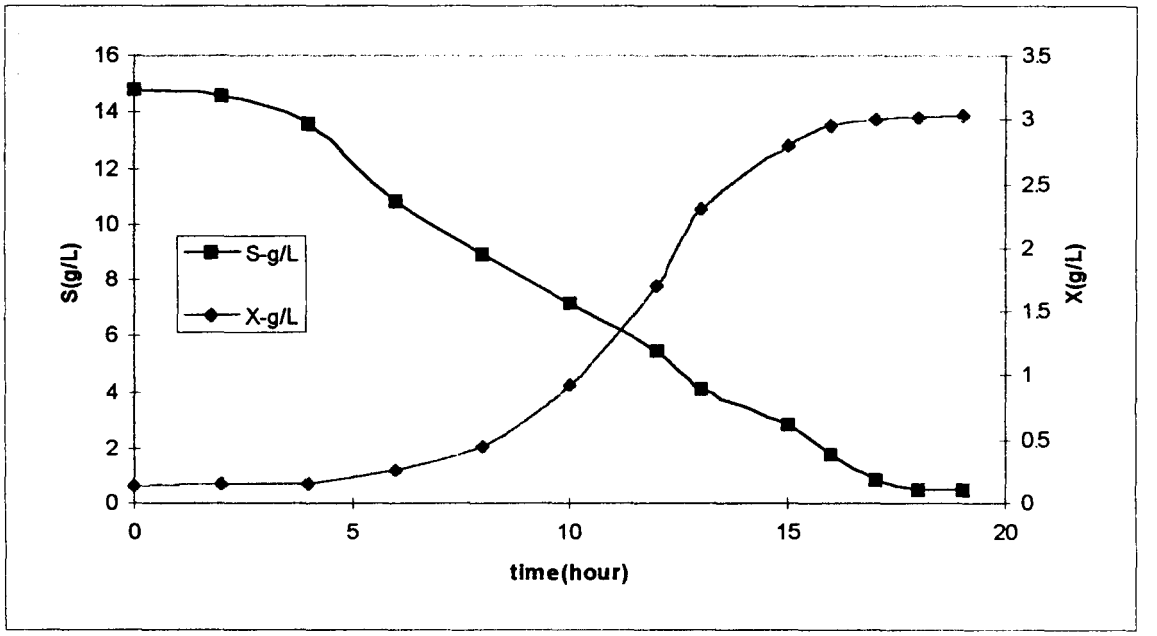


FIGURE 4.46. Time profiles of cell mass and residual glucose concentration for *E.coli*M5248 cells in minimal medium (S_0 = 15 g/L)

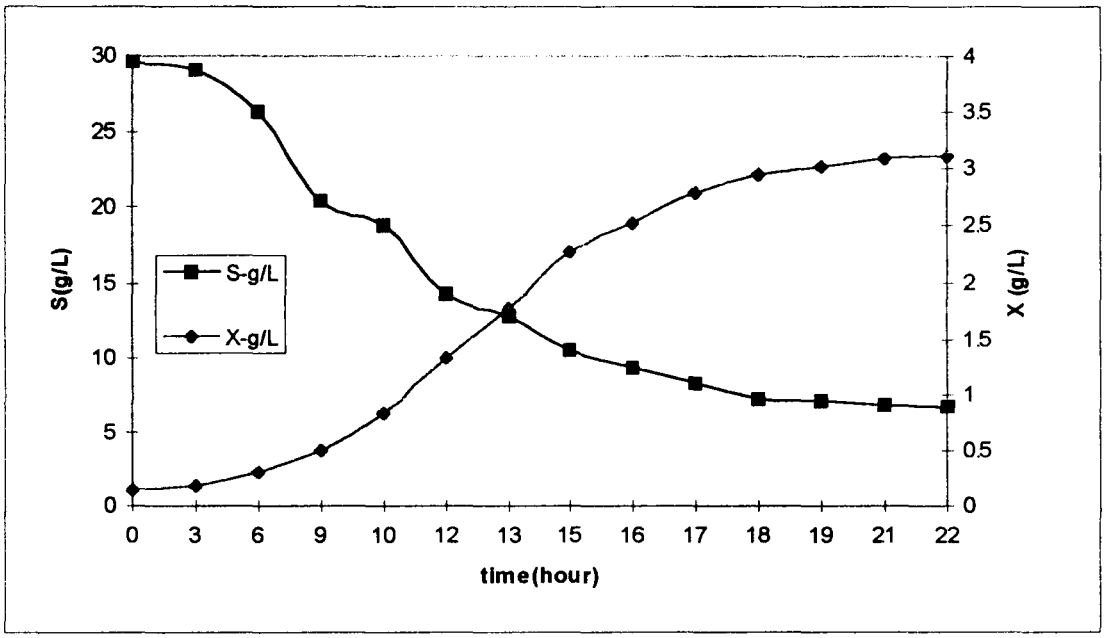


FIGURE 4.47. Time profiles of cell mass and residual glucose concentration for *E.coli*M5248 cells in minimal medium (S_0 = 30 g/L)

The specific growth rates were determined from the slope of the exponential phase of ln x versus time plots using linear regression (Table 4.13). These specific growth rates are also plotted versus initial glucose concentration for both recombinant strains (Figure 4.48). The specific growth rates (0.186 hr^{-1} for *E.coli* M5248 and 0.192 hr^{-1} for *E.coli* 294) obtained in cultures with 1g/L initial glucose concentrations are at their lowest values because of the insufficient amount of carbon sources in the growth media. As illustrated in Figure 4.48, the specific growth rate increases with increasing substrate concentration up to 10 g/L, remains almost constant at a maximum value between 10 - 15 g/L initial glucose concentration for both recombinant strains. The lowest μ_m value is obtained at 30 g/L initial glucose concentration indicating a possible substrate inhibition effect.

The results obtained show that both recombinant strains have similar growth behaviour with respect to different initial glucose concentrations under the same fermentation conditions. Again, the inhibitory effect of glucose on growth when supplied at high concentrations (>15 g/lt) was observed in both recombinant *E.coli* M5248 and *E.coli* 294 cell cultures.

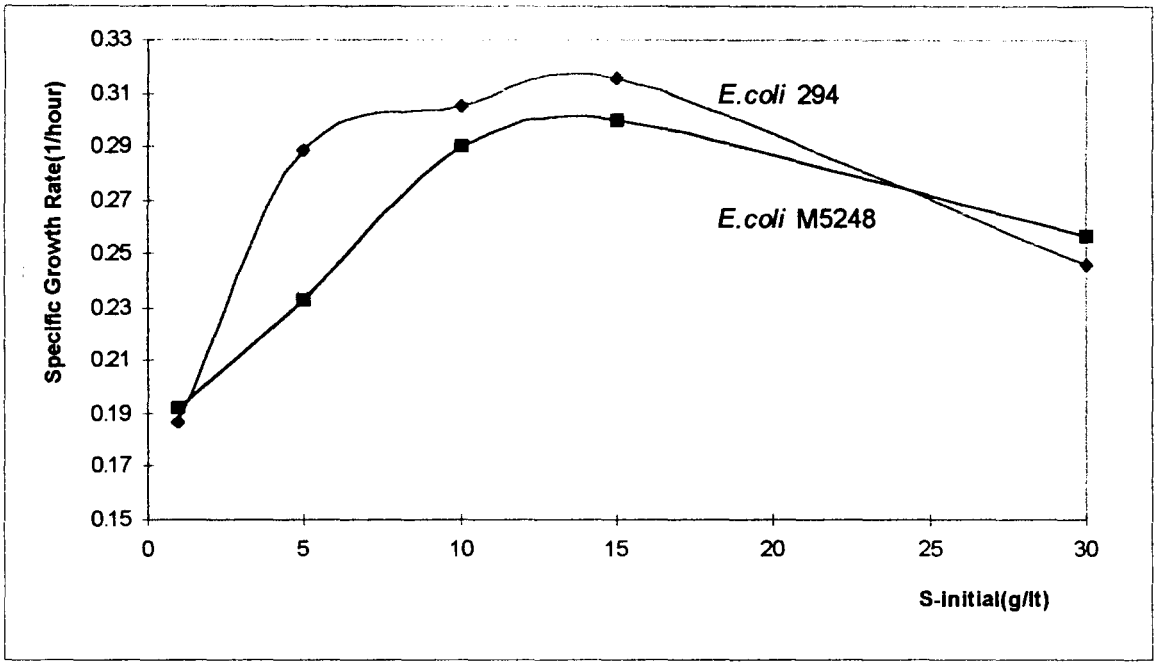


FIGURE 4.48. Maximum specific growth rates at different glucose concentrations for the recombinant strains *E.coli* M5248 and *E.coli* 294.

Samples taken from the fermentation media at regular time intervals were also analysed for their *Eco*RI enzyme activity after distrupting the cell membrane by the addition of lysosyme (Section 3.2). Time profiles of the *Eco*RI total activity are shown for both of the recombinant *E.coli* strains in Figures 4.49-4.53.

The *Eco*RI activity was not detectable until the mid exponential phase. The activities began to increase slightly at this stage, and a sudden increase was observed following the induction. The use of a different medium did not affect the optimum induction conditions; only the activities obtained at the end of the induction period were lower than the ones obtained from cultures grown in rich medium. Since enzyme production is directly related to the biomass formed, it increases with the same ratio as in the increase of biomass. In the fermentation medium with an initial glucose concentration of 1g/L, total activity and the biomass levels reached were 8,337,500 U and 1.4g/L for the *E.coli* 294 strain and 500,250 U and 1.36g/L for *E.coli* M5248 strain. At the end of the fermentation with 5g/L initial glucose concentration, both the total activity and the biomass were doubled in the case of the *E.coli* 294 strain while the total activity was increased by a factor of 2.3 and the

biomass by a factor of 1.9 in the case of the *E.coli* M5248 strain. The same trend was observed for both strains during fermentations with 10 and 15g/L initial glucose concentration.

The optimum total activity values obtained from the cell cultures grown in minimal medium with 15g/L initial glucose concentration were $5x10^7$ U and $1.67x10^6$ U from the *E.coli* 294 and *E.coli* M5248 cell cultures respectively. The activity of *Eco*RI endonuclease was clearly higher in the case of *E.coli* 294 cells as mentioned in Section 4.3 for the rich medium fermentations. It was also mentioned previously that the high initial glucose concentration of 30 g/L has an inhibition effect on the growth. The decrease in total activity of the enzyme due to inhibited biomass formation is observed in Figure 4.54.

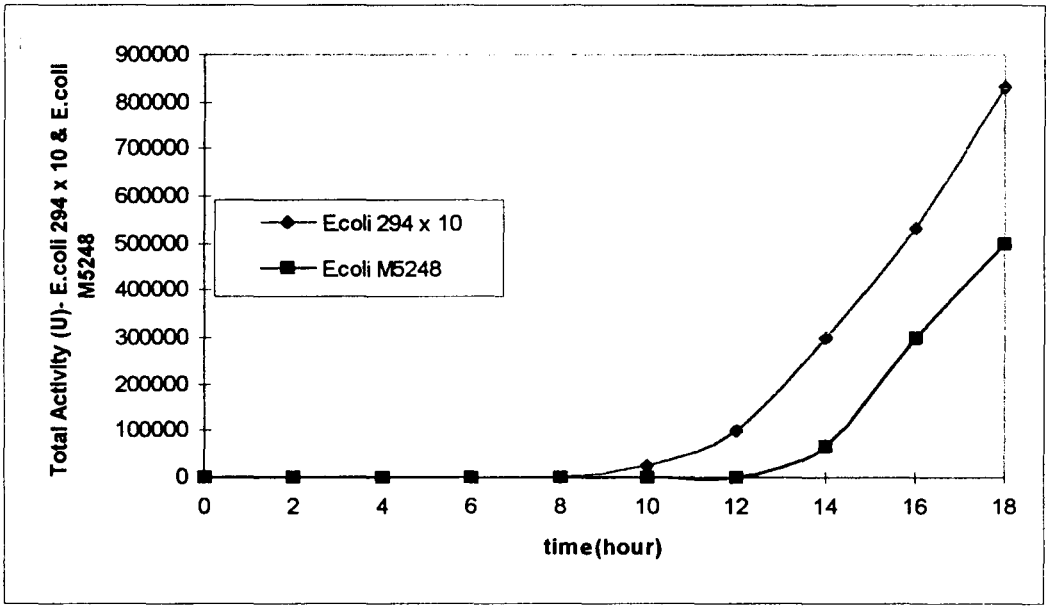


FIGURE 4.49. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 1g/L)

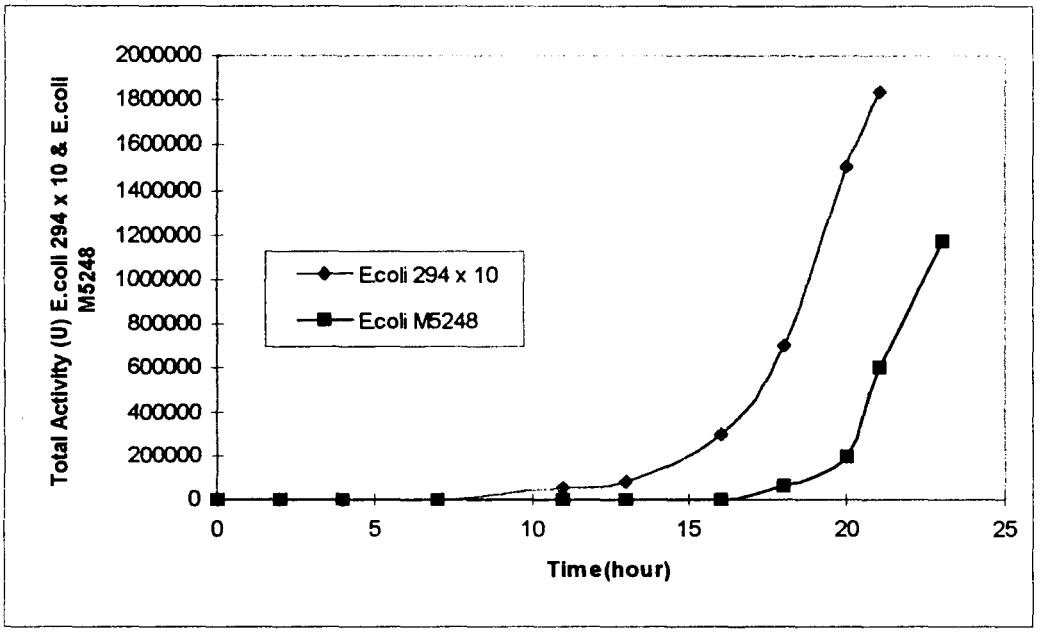


FIGURE 4.50. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 5g/L)

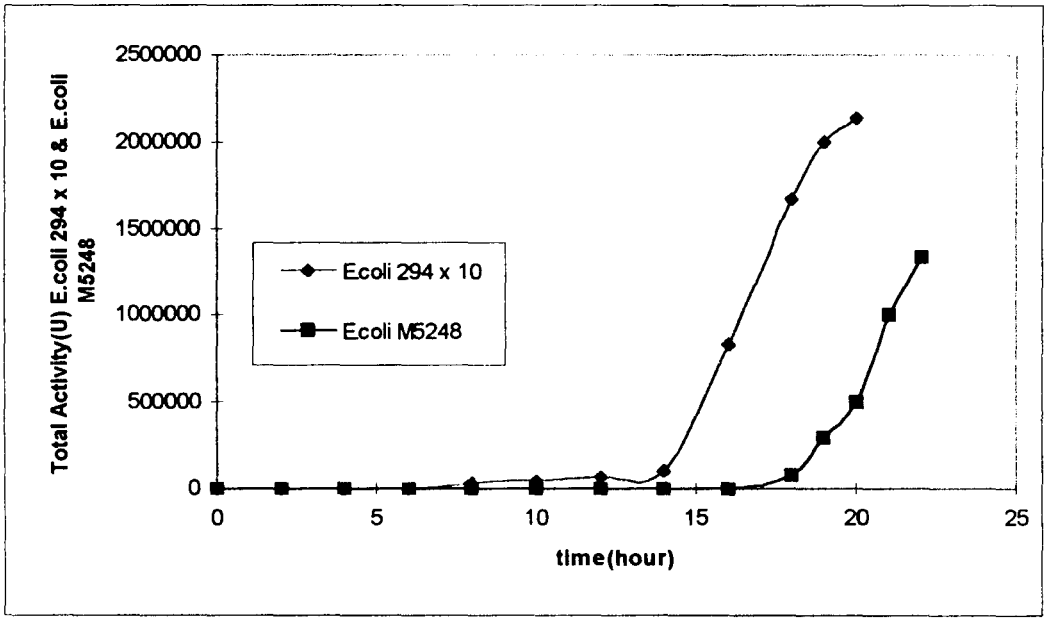


FIGURE 4.51. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 10g/L)

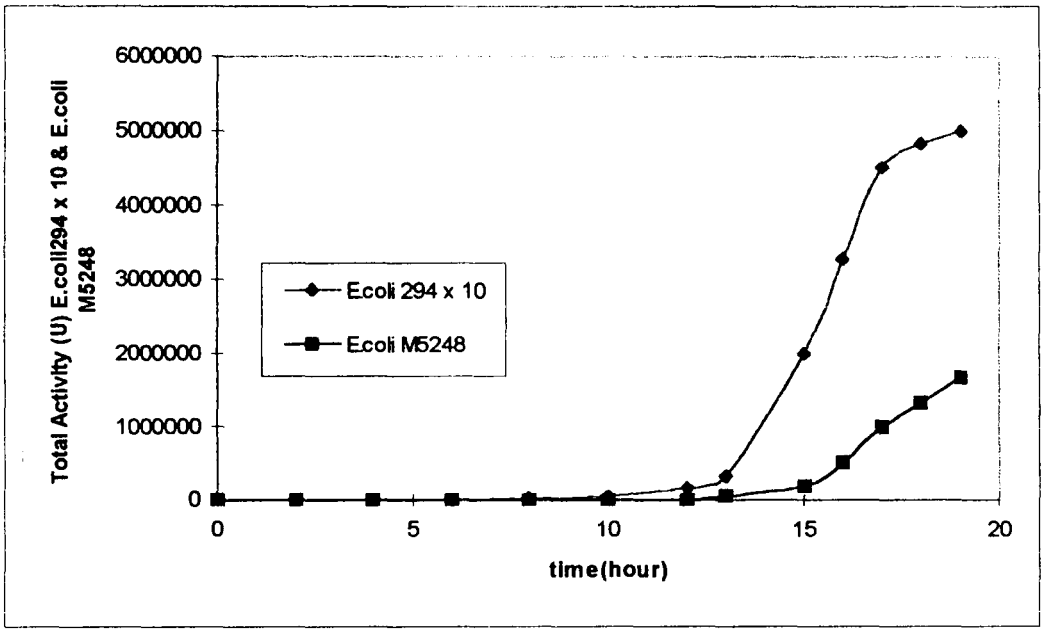


FIGURE 4.52. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 15g/L)

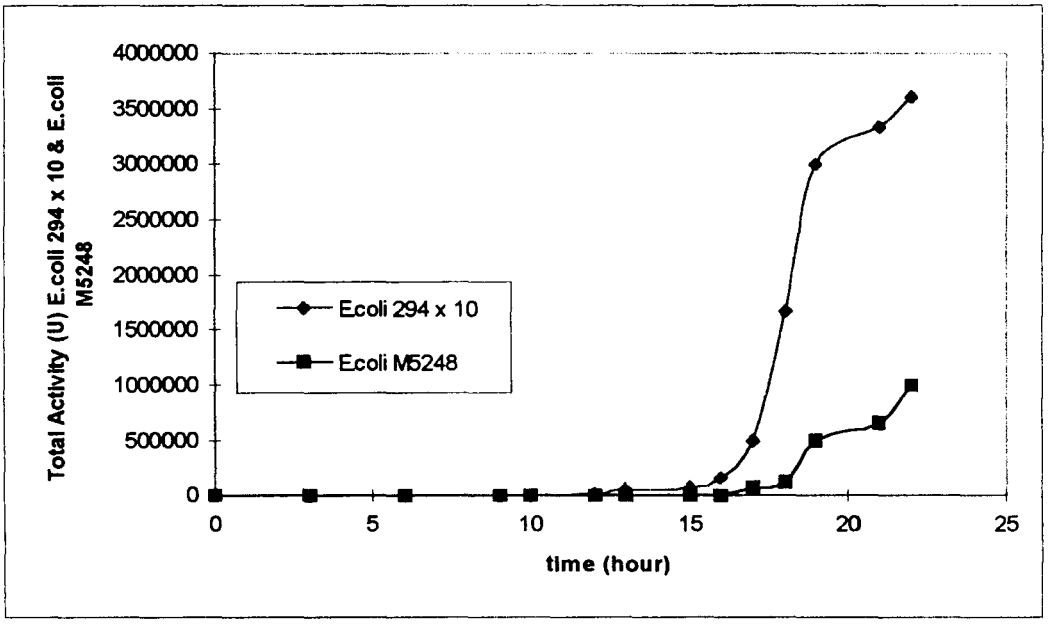


FIGURE 4.53. Time profile of the total activity of *Eco*RI endonuclease in *E.coli* M5248 and *E.coli* 294 cell cultures grown in minimal medium (S_0 = 30g/L)

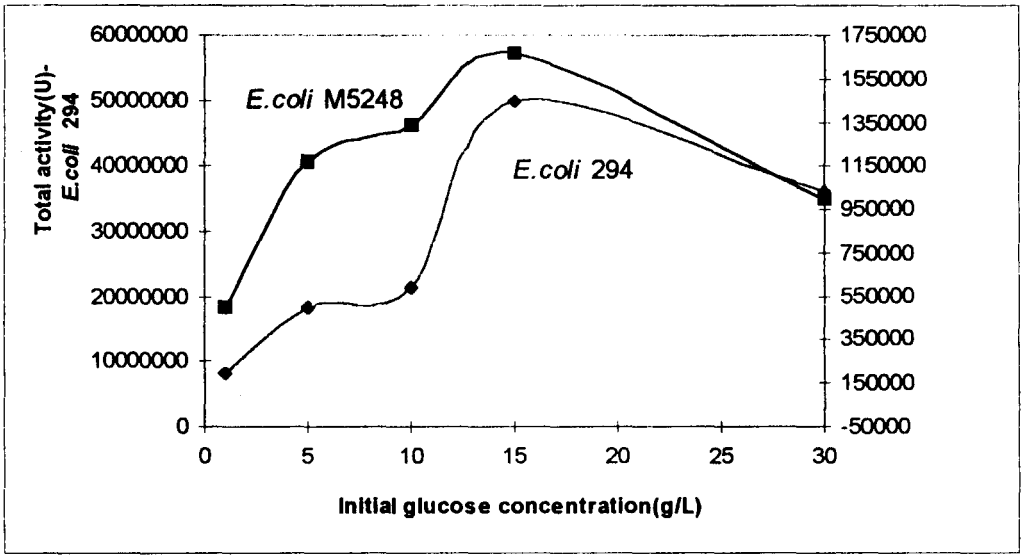


FIGURE 4.54. The effect of initial glucose concentration on the total activity of the *Eco*RI endonuclease for both of the recombinant *E.coli* strains.

4.6 The Effect of Glucose Concentration on the Intracellular Components of Recombinant *E.coli* 294 Strain

The recombinant strain giving the higher productivity of *Eco*RI endonuclease, i.e. *E.coli* 294 cells containing plasmid pPG430, were chosen as the recombinant system to investigate the change in the composition of the intracellular components during fermentation. In each experiment conducted with recombinant *E.coli* 294 strain by using different initial glucose concentrations (Section 4.5), changes in the RNA concentration, plasmid content and the total protein amount were also determined with respect to time.

In section 4.5, it was concluded that 30 g/L initial glucose concentration led to substrate inhibition and decreased cell growth. In addition to the experiments performed by using five different initial glucose concentrations, an additional fermentation experiment with 20 g/L initial glucose concentration was performed by using recombinant *E.coli* 294 strain to observe if there is a glucose inhibition effect between 15g/L-30g/L glucose.

The biomass growth and glucose consumption at 20g/L initial glucose concentration are given in Figure 4.55. In this fermentation experiment, around 2 g/L glucose remained unconsumed. The maximum specific growth rate calculated from logarithmic plot of biomass versus time, $\mu_m = 0.309\ hr^{-1}$, was found to be in the expected region of Figure 4.48, showing almost no inhibition effect by the glucose.

Shake flask fermentations with six different glucose concentrations were conducted as explained in Section 4.5, and the samples taken from the fermentation media at regular time intervals were analysed for their RNA, plasmid and total protein content after distrupting the cell membrane by the addition of lysosyme (Section 3.2.7.1).

The plasmid content was determined spectrophotometrically after isolation from the other components in the cell by the method described in Section 3.3.6. The optical density measurements were taken both at 260 and 280 nm to follow the degree of protein or RNA contaminations. The OD_{260} over OD_{280} ratio as 1.8 is the reference value for a sample of high purity, since this ratio shows the existence of protein contamination above 2.0, and RNA contamination below 1.8 (Coppela, et al., 1986).

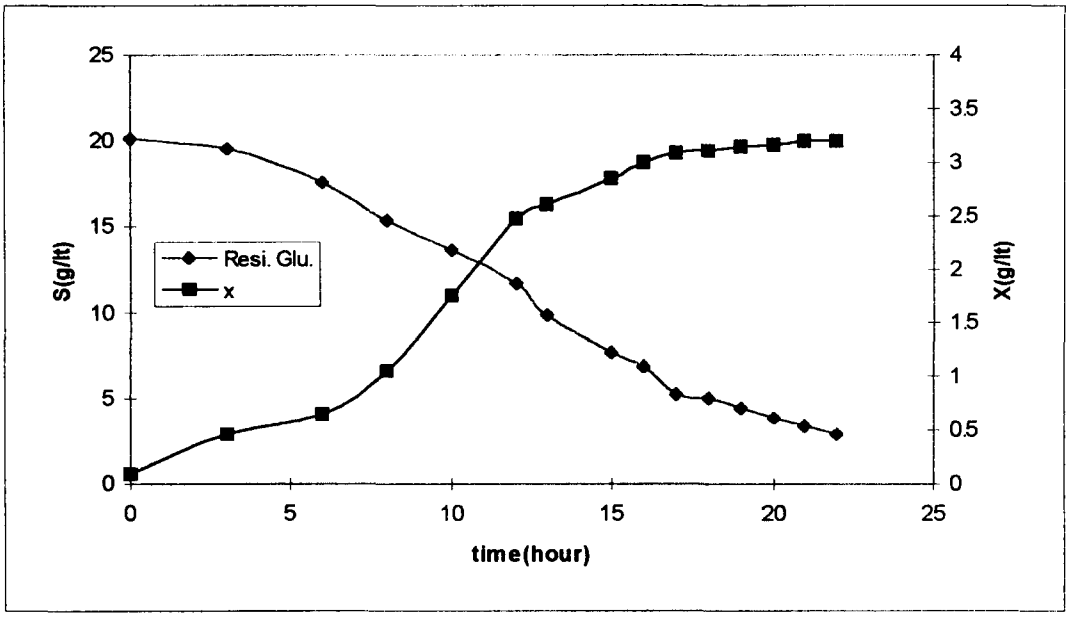


FIGURE 4.55. Time profile of cell mass and residual glucose concentration for *E.coli* 294 cells in Minimal medium with initial glucose concentration of 20g/L.

The RNA content was determined by Orcinol Method as described in Section 3.3.5 (Nielsen,1991a).The determination of the RNA content is especially important, since RNA constitutes a considerable part of the protein synthesizing system. It is directly proportional to the biomass concentration. In structured models, it is one of the key variables during fermentation (Nielsen,1991a).

The time profiles of the RNA and plasmid contents are shown in Figures 4.56-4.61 for the fermentations performed at different initial glucose concentrations. The variations of RNA concentration with time shows similar trends with growth curves of the cells. RNA concentrations varied from $2.34x10^{-4}$ to $7.96x10^{-4}$ g/mL for 1-30 g/L initial glucose concentrations and were found to be directly proportional to the biomass. Since the major component in the active part of the cell is RNA, the amount of RNA increases with time and then enters the stationary phase in each of the fermentation experiments. A similar increase in RNA with growth rate was reported also for many other organisms. Nielsen et.al(1991b) suggested that the energy source (glucose) is a determining factor for the protein synthesizing machinery active size of the machinery.

Figure 4.56-4.61 indicate that the plasmid DNA content of the cells increases at the beginning of the exponential phase, followed by a slight decrease in the amount of DNA when the concentration of RNA begins to increase. The decrease in the DNA concentrations was observed at about the same period when the activity of the *Eco*RI endonuclease was detected (Figure 4.49-4.53), and the amount of plasmid DNA becomes constant after the cells have entered their late exponential phase. These time profiles are consistent with other results reported on the variation of plasmid content in fermentation media (Rehm and Reed, 1991). The final amount of plasmid content varied from $1.14x10^{-4}$ to $2.98x10^{-4}$ gP/gcell in recombinant cultures grown in minimal medium with 1-30g/L initial glucose concentration, respectively.

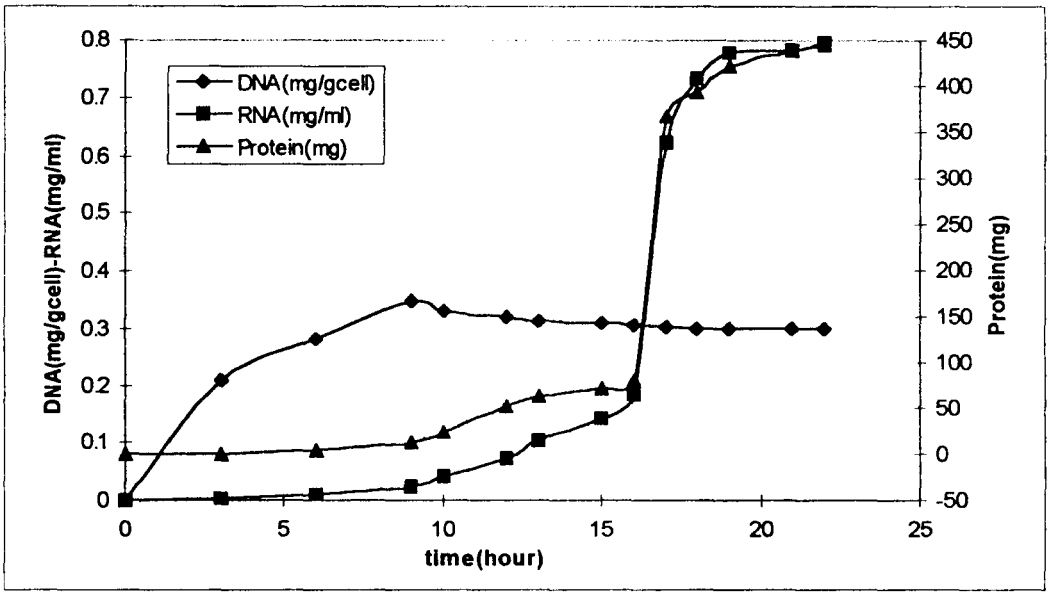


FIGURE 4.56. Time profiles of RNA, plasmid DNA content amd total protein content for 30g/L initial glucose concentration (*E.coli* 294)

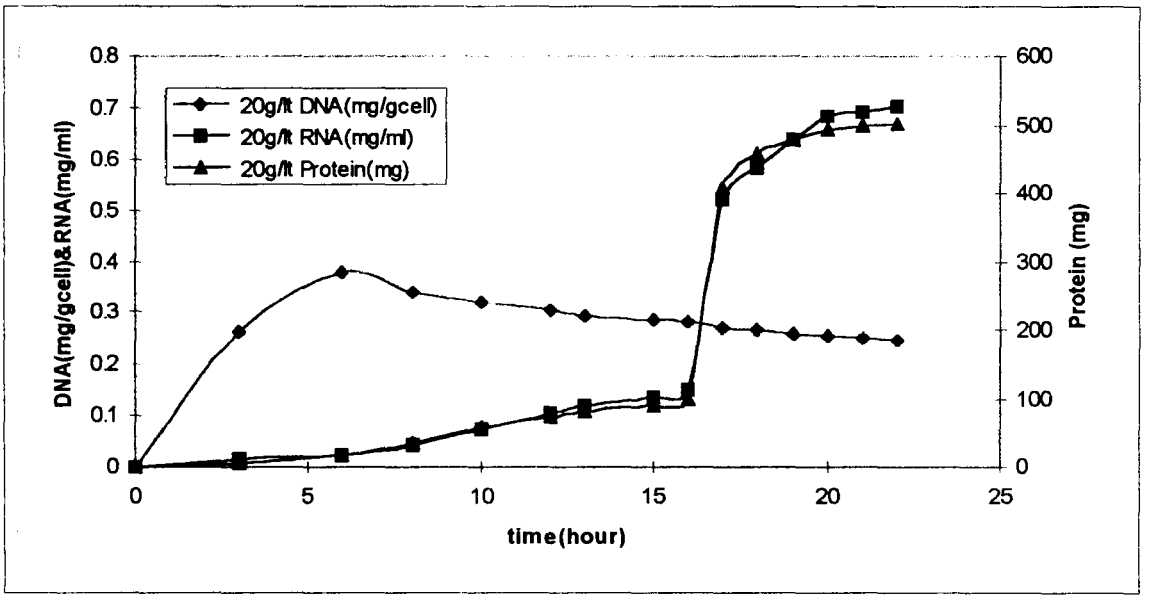


FIGURE 4.57. Time profiles of RNA, plasmid DNA content amd total protein content for 20g/L initial glucose concentration (*E.coli* 294)

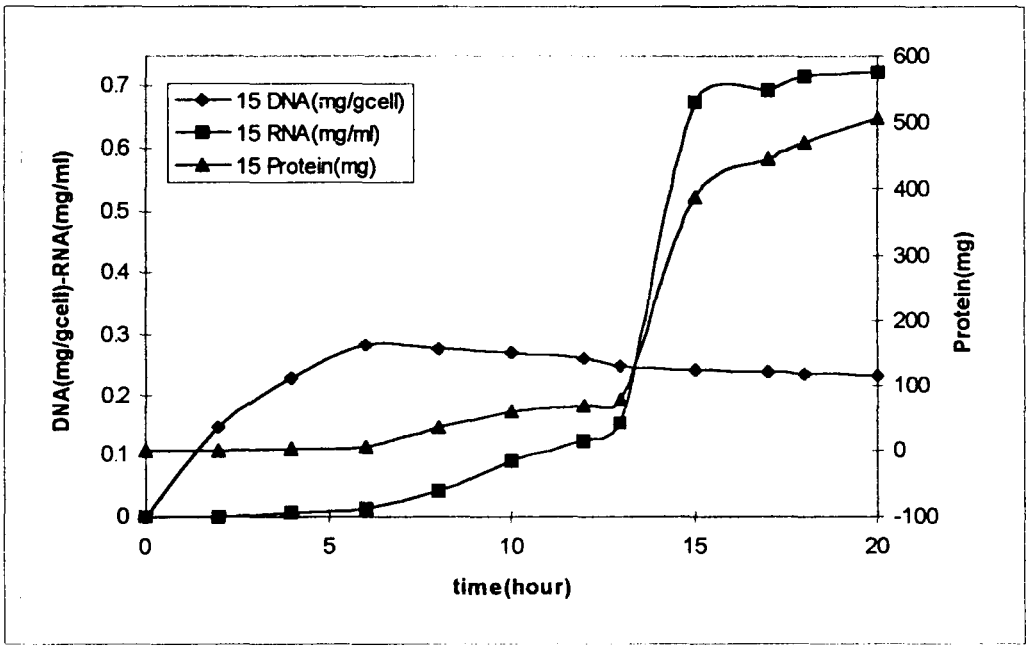


FIGURE 4.58. Time profiles of RNA, plasmid DNA content amd total protein content for 15g/L initial glucose concentration (*E.coli* 294)

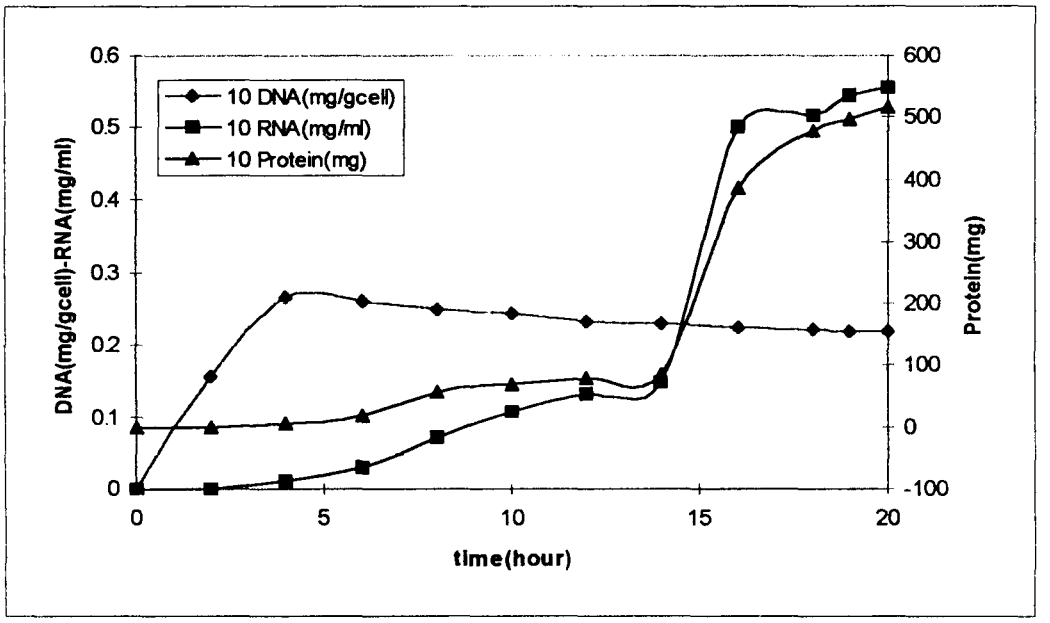


FIGURE 4.59. Time profiles of RNA, plasmid DNA content amd total protein content for 10g/L initial glucose concentration (*E.coli* 294)

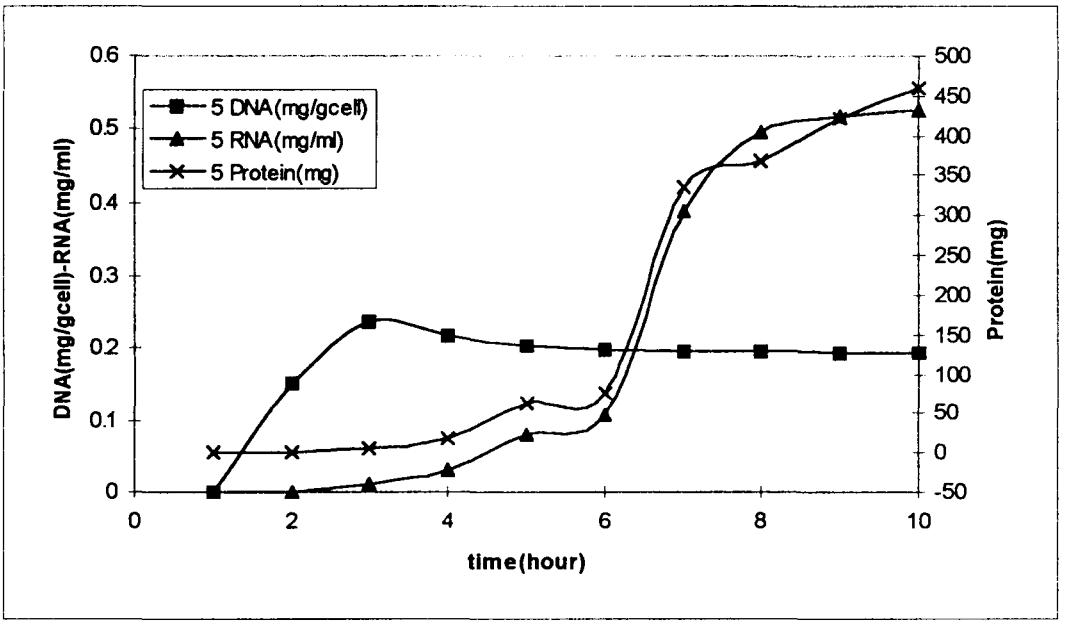


FIGURE 4.60. Time profiles of RNA, plasmid DNA content amd total protein content for 5g/L initial glucose concentration (*E.coli* 294)

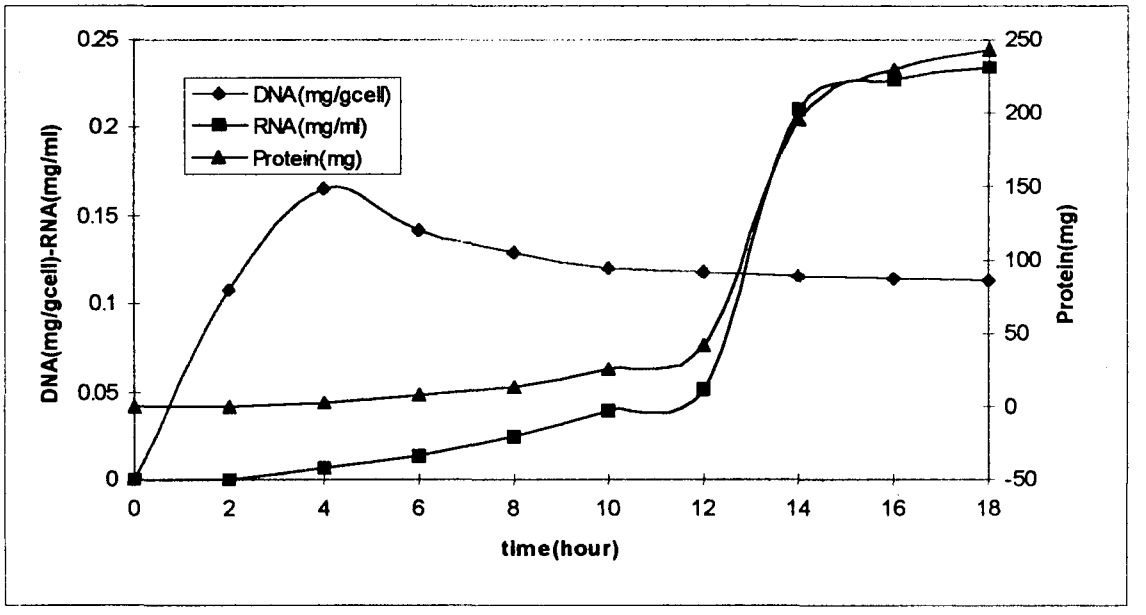


FIGURE 4.61. Time profiles of RNA, plasmid DNA content amd total protein content for 1g/L initial glucose concentration (*E.coli* 294)

Time profile of the *Eco*RI endonuclease activity in the *E.coli* 294 culture grown in minimal medium with 20 g/L initial glucose concentration is shown in Figure 4.62. The total activity was found to be $4.4x10^7$ U which is greater than the activity found in the cultures grown in 30g/L initial glucose concentration but smaller than those at 15 g/L initial glucose concentration. This supports the result which was obtained at the end of Section 4.5 as the inhibition of glucose shows its effect after the 15g/L initial glucose concentration.

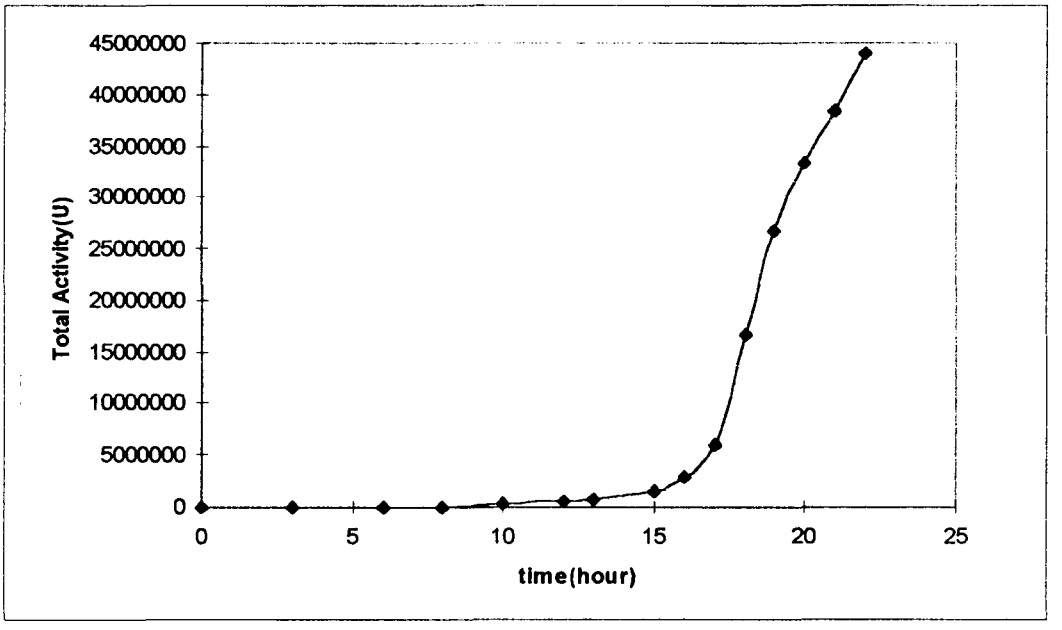


FIGURE 4.62. Time profile of total *Eco*RI activity in recombinant *E.coli* 294 cells grown in minimal medium (So=20g/L).

4.7. Investigation of Copy Number by Using Segregated Models

In the previous sections, the productivity of the cells was found to be directly related to the plasmid content of the recombinant cell population. Plasmid stability was determined from the relative numbers of plasmid-containing and plasmid-free cells. Plasmids concerned (pPG430 in *E.coli* 294 and pSCC2 in *E.coli* M5248) were unstable in the sense that the

fraction of the cell population containing plasmids declined from unity initially to smaller values as growth proceeded.

Plasmid copy number is also of particular importance in the use of genetically modified bacterial cultures. It is a parameter used in describing recombinant bacterial population, and the distribution of the cellular plasmid content around the mean value may be quite wide in individual cells in the population. Individual cells containing different numbers of plasmids are expected to have different probabilities of plasmid loss at division.

In this part of the study, plasmid copy number of the average cells and the plasmid number distribution at birth within the population were determined by using the mathematical models based upon a segregated, population balance approach (Seo and Bailey, 1985).

The plasmid copy number was calculated from the equation derived by Seo and Bailey (1985). They have proposed an equation for the calculation of the plasmid-bearing fraction of the population after m generations :

$$\phi_p(m) = \frac{1}{1 + \frac{\theta(1-\eta)}{(1-\theta)\cdot(2-\eta)} \cdot \left[1 - \left(\frac{2^{\alpha}}{2-t}\right)^{m}\right]} \qquad (4.14)$$

where $\eta = (2-\theta)^{1/\alpha}$ (4.15) and $\theta = 2^{1-CN}$ (4.16)

The probability of plasmid loss, θ, has been first proposed by Koizumi et al. (1984) in the investigation of the effect of plasmid instability on the copy number. Ataai and Shuler (1987), Ryan, W. and J. S. Parukelar (1990) and Shu and Shuler (1992) have used the equations 4.14 - 4.16 to find the relation between plasmid copy number and the plasmid instability. They have expressed the fraction of plasmid harboring cells ϕ p as a function of two parameters: the single cell growth rate ratio, α which may be controlled by adjustment of the growth environment, and the probability of plasmid loss, θ, which is

characterized by plasmid genetic functions. Parameter $\theta = 2^{1-CN}$ was inserted into the equation for calculating the plasmid copy number, CN. The program given in Appendix ..... was used to calculate the copy numbers for given values of α, fraction of the plasmid containing cells and the number of generations m. The copy numbers were first investigated by setting the single cell growth rate ratio, α equal to unity, since the growth rates of the plasmid-containing and plasmid-free cells on the single cell level are identical. Later on, the α value has been changed, because the fraction of the plasmid-containing cell population decreases with the number of generations due to plasmid instability. The fraction of the plasmid-containing cells were calculated from the experimental data presented in Section 4.4 for both strains.

The plasmid copy numbers which were estimated from the model equation show different patterns for different single-cell growth rate ratio (α). When α is equal to unity, copy numbers vary from 4-5 to 9-10 for both strains *E.coli* 294 and *E.coli* M5248 as the number of generations increases. The model predicts an increasing copy number with increasing number of generations but the number is quite stable after a certain generation.

In *E.coli* 294 (pPG430) strain, the plasmid stability was between 90-100% until the induction of the cells and during the induction period the stability decreased down to 85%. The copy number of the pPG430 was found to be five until induction with IPTG, then this value was found to increase to 14 (Figure 4.63). In *E.coli* M5248 (pSCC2), the copy number of the plasmid pSCC2 was found to be 10 at the late exponential phase of growth. Following the induction by T-shift, the copy number of the plasmid pSCC2 reached 38, although plasmid stability decreased down to 50% at the end of the induction period (Figure 4.64). Plasmid pSCC2 had an increased level of productivity, because the increase in the copy number following the T-shift compensated the decrease in the plasmid stability. Brantl and Behnke (1992) have investigated the change in plasmid copy number of plasmid pRS39 by varying the concentration of the inducer IPTG in *Bacillus subtilis* cells. They have found that the copy number of pRS39 increased in parallel to the amount of inducer added. The plasmid copy number was increased to 50 from 5 as they have induced the cells by 100 μM IPTG. Similar result was obtained by Schendel et al.(1989) in *E.coli* MZ9387 cells containg plasmid pIU106. They have reported approximately six fold increase following the

inducion by IPTG. Another study was perfomed by Du Poet et al. (1987) who have studied the relationship between plasmid copy number and the xylane activity in *E.coli* JM105 cells (pKK223-200). They have induced the cells by IPTG and reached to approximately five fold increase in the copy number. The relation between plasmid copy number and the thermal induction was undertaken in *E.coli* CSH22 (pVH106 / 172) by Betenbaugh et al. (1989). The plasmid content in terms of copies per cell was given as 170 after the temperature-shift whereas the copy number was given as 5 in the cells which were not exposed to T-shift. Vila et al.(1994) have investigated the copy number of plasmid pJC046 in *E.coli* MC1061 cells under batch fermentation conditions. Following the thermal induction, they have reported a two fold increase in the plasmid copy number.

The plasmid copy numbers in Figures 4.63 and 4.64 may not coincide with the actual plasmid copy numbers that need to be determined experimentally. However, the calculated CN values indicate the increase in the copy numbers of both plasmids after induction.

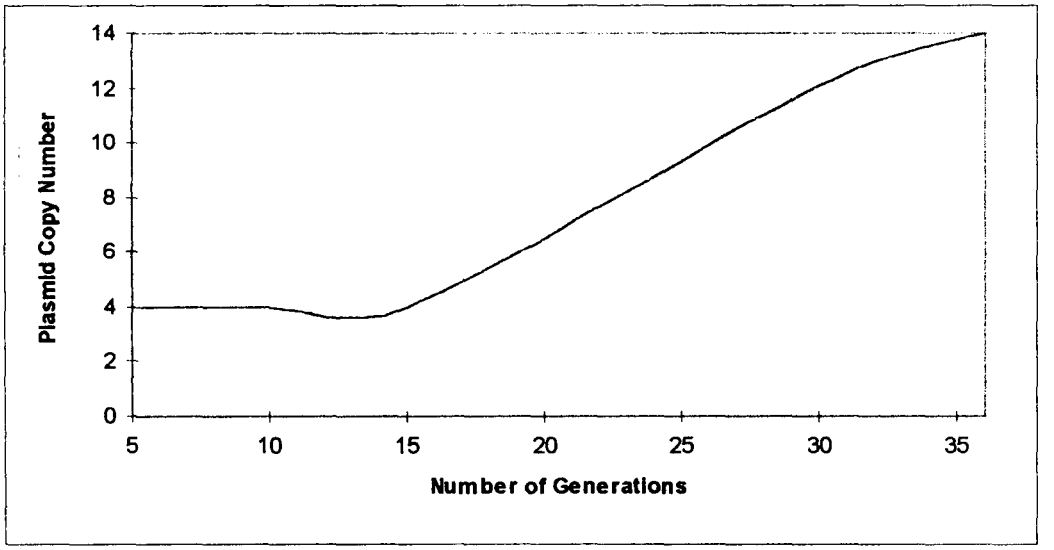


FIGURE 4.63. The effect of the number of generations on the plasmid copy number (*E.coli* 294-pPG430)

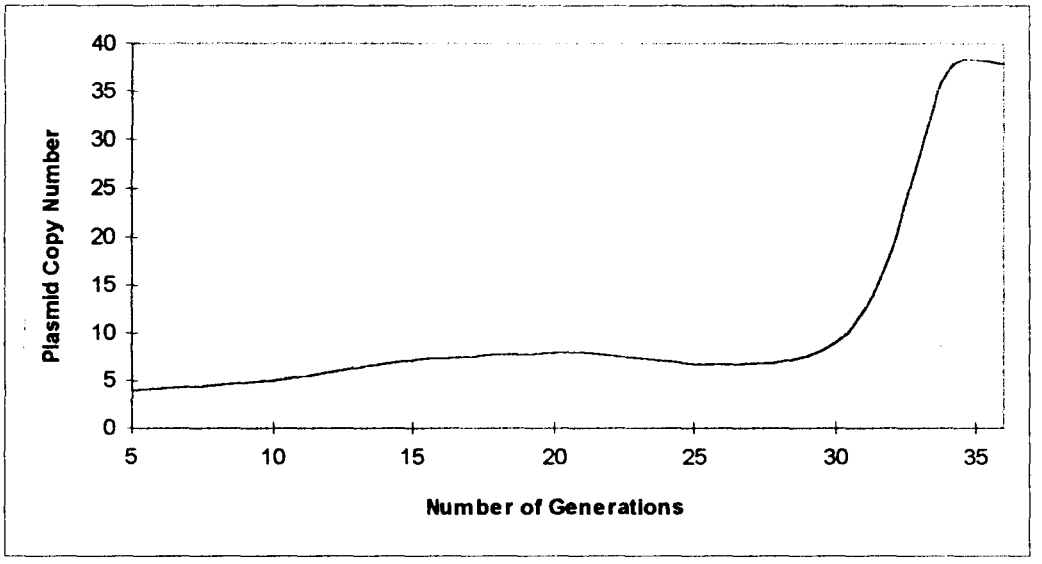


FIGURE 4.64. The effect of the number of generations on the plasmid copy number (*E.coli* M5248 - pSCC2)

4.8 Growth Kinetics Using Unstructured Models

In unstructured models, the microorganism is considered as a single reacting species with fixed chemical composition, and therefore cell mass concentration is used to characterize the biophase. Other major components of the fermentation system are the limiting substrate and excreted products. Control models for the routine operation of industrial fermentations are often based on simple unstructured models which are sufficient in approximating system behaviour for many technological purposes.

Cell growth may be inhibited in various ways, i.e. by the presence of chemical inhibitors or by particular levels of product and/or substrate concentrations. In unstructured models substrate utilization, product formation and the effect of product or substrate inhibition on cell growth are considered. In the present study, the inhibitory effect (if any) of product can not be taken into account in unstructured modelling, since the *Eco*RI endonuclease produced is an intracellular product and is not present in the medium.

The inhibitory effect of substrate has been experimentally observed at relatively high initial substrate concentration, i.e. above 15g/L glucose in minimal medium.

The experimental data obtained from the batch fermentation of recombinant *E.coli* 294 and *E.coli* M5248 cells at five different initial glucose levels (1, 5, 10, 15, 30 g/L) were used to model the biomass production by these strains. Four different mathematical models that have been reported in the literature for describing the effects of substrate inhibition on cell growth were tested to find the model which represents the experimental data obtained.

In Model 1 given by Equation (4.17), biomass formation is based on the generalized nonlinear equation proposed by Luong (1987) for correlating substrate inhibition.

Model 1

$$\mu = \frac{\mu_m \cdot S}{Ks + S} \cdot \left[1 - \frac{S}{Sm}\right]^n \qquad (4.17)$$

The four-parameter model is of the generalized Monod type and accounts for both substrate stimulation at low S_0 and substrate inhibiton at high S_0. The model can be used to predict S_m, the maxiumum substrate concentration above which growth is completely inhibited.

Model 2, given by Equation (4.18) below has been proposed by Edwards (1970) and is a modified version of the equation that was initially derived by Teissier (1942):

Model 2

$$\mu = \mu_m \left(e^{-\frac{S}{K_i}} - e^{-\frac{S}{K_s}} \right) \qquad (4.18)$$

In this three-parameter model, the assumption of diffusion-controlled substrate supply has been combined with substrate limitation at high and inhibitory concentrations. Although substrate inihibition is included, the model is not suitable for predicting the substrate concentration level, S_m, at which growth stops.

The third mathematical representation, Model 3 given by Equation (4.19), has also been proposed by Edwards (1970):

Model 3 $$\mu = \frac{\mu_m \cdot S}{K_S + S + \left(\frac{S^2}{K_i}\right) \cdot \left(1 + \frac{S}{K_P}\right)} \qquad (4.19)$$

The equation has four parameters and is a modified form of the Haldane equation used for describing enzyme inihibition by the formation of inactive complexes of the enzyme with more than one substrate molecule. The model implies indefinite cell growth and, again, the maximum substrate concentration, S_m, above which cells cease growing is not predicted.

Model 4 given by Equation (4.20) was initially proposed by Aiba et al (1968) for product inhibition and later borrowed by Edwards (1970) to correlate substrate inhibiton.

Model 4 $$\mu = \frac{\mu_m \cdot S}{K_S + S} \cdot e^{\frac{-S}{K_i}} \qquad (4.20)$$

This three-parameter model represents exponential kinetics and does not predict the S_m value for complete inhibition.

4.8.1 Overall Yield Factors

Overall yield factors are sometimes used to characterize system performance and are based on the total biomass or total product produced during the fermentation period. Before testing the applicability of the model equations to the fermentation data, the relations between biomass, substrate and product concentrations were investigated in terms of the overall yield factors. The overall yield factors for biomass formation per substrate consumption, Yx/s (g cells formed per g substrate consumed), and the product formation per biomass production Yp/x (g product formed per g cells formed) are given in Table 4.14 for both strains and plotted in Figure 4.65-66. For the *E.coli* 294 strain the Yx/s value remains almost constant after 15g/L initial glucose concentration which implies that the

increase in the initial substrate concentration does not cause an increase in the relative biomass formation (Figure 4.65). The Yp/x value shows that product formation is maximum at 15g/L initial glucose concentration (Figure 4.66) and at higher concentrations of the substrates, product formation starts to decrease. This indicates an inhibitory effect of the substrate on product formation above 15g/L initial glucose concentration. For the *E.coli* M5248 strain, on the other hand, Yp/x is maximum at 10 g/L and there is substrate inhibition on product formation after 10g/L initial glucose concentration (Figure 4.66). Since there is a change in Yx/s after the same substrate concentration, the inhibitory effect of the subtrate on cell growth is more important than in the case of *E.coli* 294.

TABLE 4.14. The overall yield factors Yx/s and Yp/x for the recombinant *E.coli* 294 and *E.coli* M5248 strains

Strains	**Initial Glucose Concentration (g/L)**	**Y x/s (g cells/g substrate)**	**Y p/x (g protein/ g cells)**
***E.coli* 294**	1	1.428	0.292
	5	0.55	0.296
	10	0.308	0.321
	15	0.219	0.759
	30	0.218	0.419
***E.coli* M5248**	1	1.193	0.196
	5	0.481	0.231
	10	0.272	0.541
	15	0.202	0.274
	30	0.129	0.107

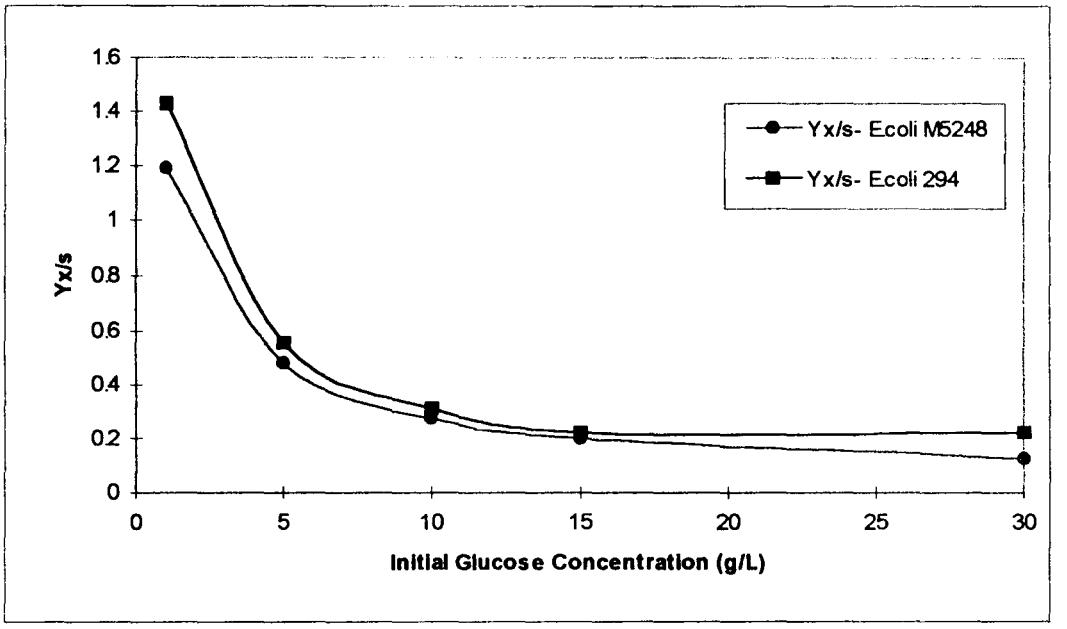


FIGURE 4.65. The overall yield factor Yx/s as a function of initial glucose concentration for the recombinant strains *E.coli* 294 and , *E.coli* M5248.

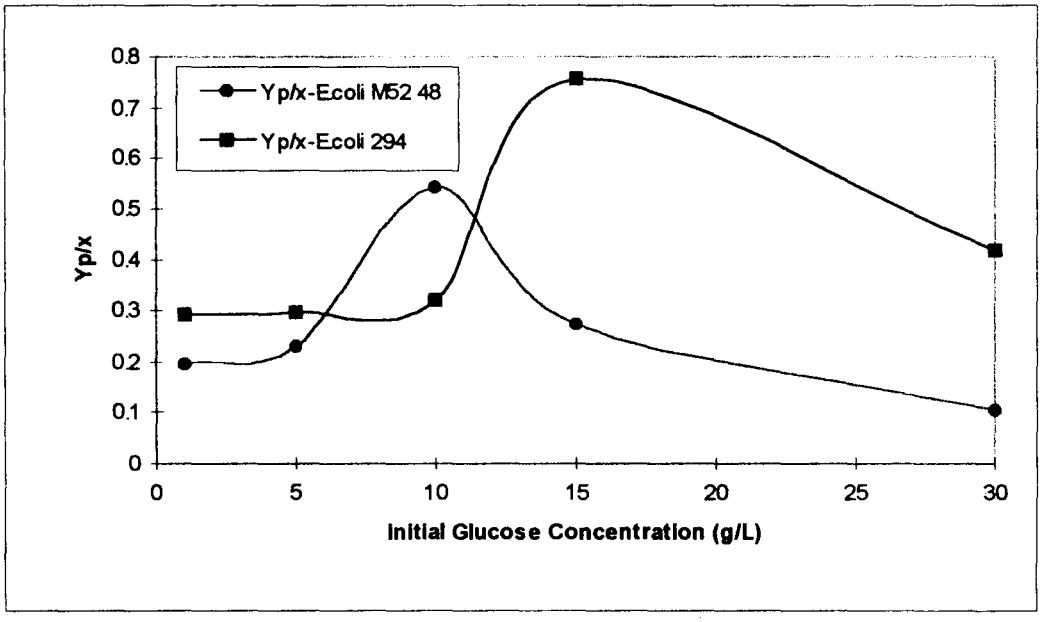


FIGURE 4.66. The overall yield factors Yp/x as a function of initial glucose concentration S_0 for the recombinant strains *E.coli* 294 and *E.coli* M5248

4.8.2 Maximum Specific Growth Rates

In the second part of the study, the maximum specific growth rates that have been obtained for each initial glucose concentration S_0 by linear regression of the ln x versus time data in the exponential phase (Section 4.5, Table 4.13) were fitted to the model Equations 4.17, 4.18., 4.19 and 4.20 by the Levenberg-Marquardt nonlinear least squares regression method for the estimation of the model parameters.

The comparison of the performance curves obtained from the model equations with the experimental data is illustrated in Figure 4.67 for the recombinant *E.coli* 294 strain. All model equations except Model 3 appear to be useful for representing the experimental data. Model 2 predicts that the specific growth rate increases with increasing S_0 from 0 to about 6 g/L followed by a gradual decrease with further increase in S_0. A somewhat similar result is obtained with Model 4. The experimental data indicate that the specific growth rate increases rapidly up to about 6 g/L, the increase continues gradually up to about 15g/L, remains the same until about 20g/L and then a decline follows. Model 1 predicts this general trend perfectly.

The performance curves obtained from Models 1, 2 and 4 are compared with the experimental data in Figure 4.68 for the recombinant *E.coli* M5248 strain. Model 3 did not give a reasonable fit predicting the substrate inhibition. Among Models 1, 2 and 4, the decline in the specific growth rate observed in the experimental data above 15 g/L is represented quite closely by the performance curve of Model 1.

The parameters of the performance curves plotted in Figure 4.67 and 4.68 are given in Table 4. 15. Model 3 was discarded from the modelling of *E.coli* 294 since the standart deviation was too high. The standart deviations of Model 2 and 4 are acceptable, and the model parameters obtained are meaningful and reasonable. Both models are not very successful in predicting the experimental observations in the higher S_0 range. The parameters and the standart deviation obtained from Model 1 show that this model is the best-fitting one although the S_m value of 30.17 g/L is somewhat low for the complete

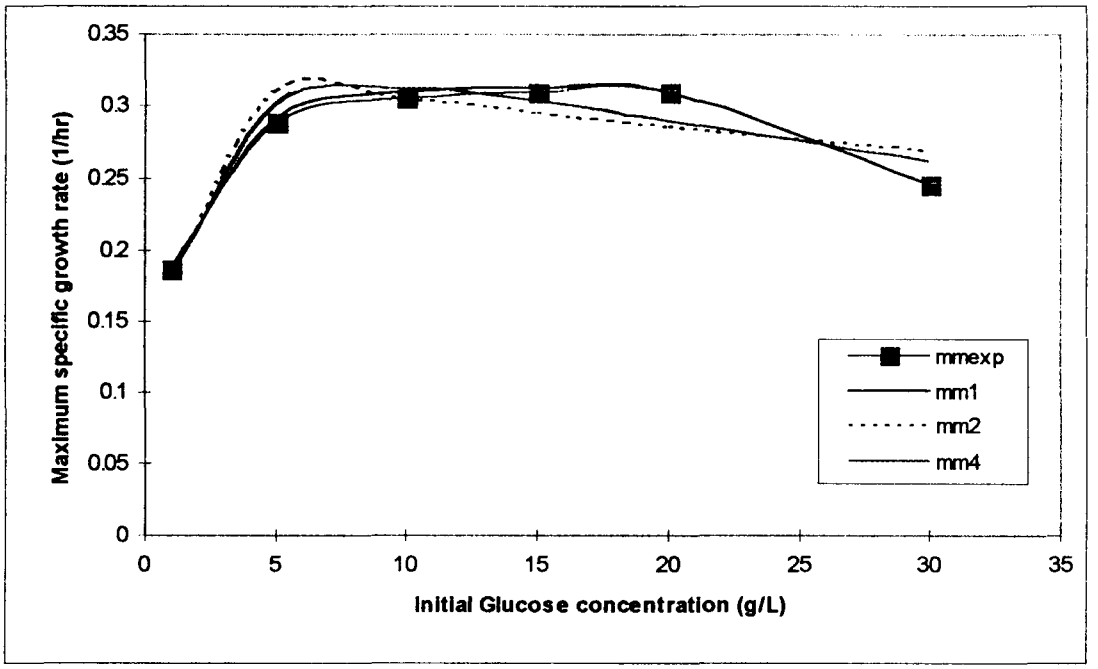


FIGURE 4.67 The effect of initial glucose concentration on the maxiumum specific growth rate for the *E.coli* 294 strain

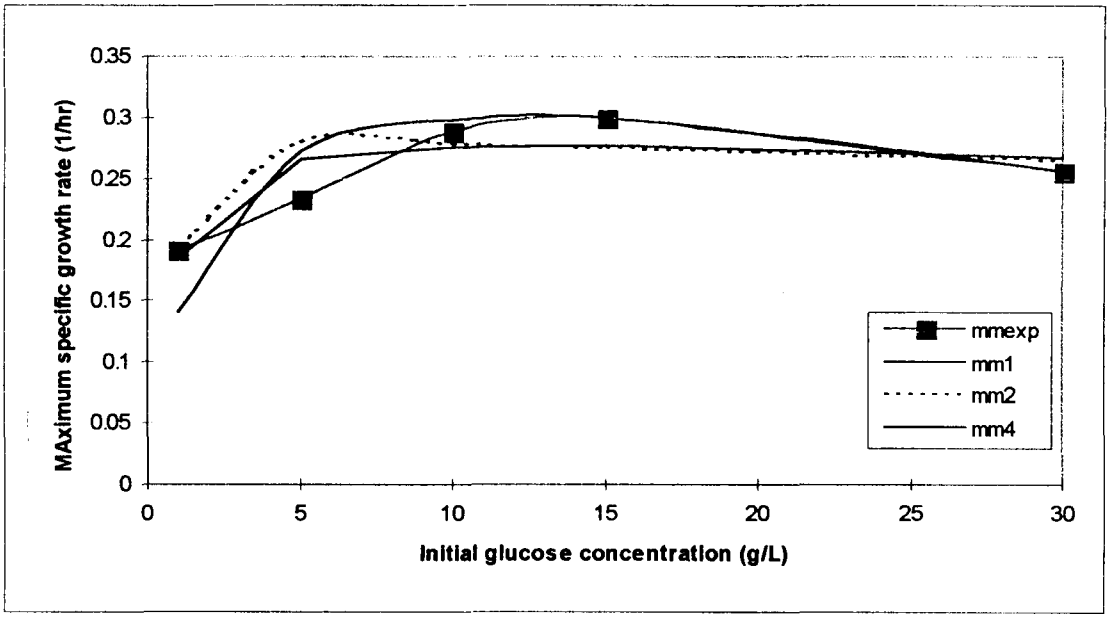


FIGURE 4.68. The effect of initial glucose concentration on the maxiumum specific growth rate for the *E.coli* M5248 strain

inhibition of cell growth. Table 4.15 shows that the standart deviations obtained from Models 1, 2 and 4 for the *E.coli* M5248 strain are close to each other. The parameter values obtained are reasonable and meaningful. The F-test for the equality of variances between these equations, i.e. ($\sigma^2_{\text{largest}}$ / $\sigma^2_{\text{smallest}}$), shows that they are not significantly different from each other in predicting the behaviour of *E.coli* M5248.

TABLE 4.15. Parameter values obtained from the modelling of maximum specific growth rate versus initial glucose concentration data

***E.coli* 294 strain**					
Model 1		**Model 2**		**Model 4**	
μ_m	0.3433	μ_m	0.3261	μ_m	0.3936
S_m	30.174	K_i	154.572	K_i	80.758
K_s	0.8328	K_s	1.19142	K_s	1.141
n	0.06				
std deviation	0.0061	**std deviation**	0.0437	**std deviation**	0.0309
f-test	1.041	**f-test**	1.118	**f-test**	1.040
***E.coli* M5248 strain**					
Model 1		**Model 2**		**Model 4**	
μ_m	0.371	μ_m	0.28547	μ_m	0.3044
S_m	40.8356	K_i	96.279	K_i	82.53
K_s	1.6112	K_s	0.887	K_s	0.642
n	0.2403				
std deviation	0.0648	**std deviation**	0.056	**std deviation**	0.044
f-test	1.235	**f-test**	1.36	**f-test**	1.155

The results obtained from the modelling of cell growth indicate that Model 1 is quite successful in predicting the growth data for both *E.coli* 294 and *E.coli* M5248. Suitable model equations need to be selected for describing substrate utilization as a function of time before Equation 4.17 can be used for simulating the x versus t and s versus

t curves obtained from the fermentations conducted by using different initial glucose concentrations.

4.9 Determination of Kinetic Parameters by Using Structured Models

Unstructured models of the types considered in Section 4.8 describe only the quantitiy of the biological phase. In situations where the chemical composition of the cell changes significantly and this change influences the kinetics, structured models have to be used. A group of potentially useful structrured models are obtained when the biophase is compartmentalized into a smaller number of key components.

In this part of the study, the two-copartment and four-compartment models proposed by Nielsen et al. (1991a, 1991b) were modeified to describe the dynamic changes in the intracellular components of *E.coli* 294 by a three compartment structured model.

The fundamental assumption is that the cell can be divided into three compartments. The A-compartment represents ribosomes, mRNA, tRNA and various activated building blocks. This compartment is the active part of the cell and its size can be estimated by measurement of RNA. The G compartment comprises genome and structural lipids and carbohydates. The P-compartment contains only the main plasmid product. In the model equations, compartment concentrations are given in g/g cells, Xi, and therefore the sum of the concentrations is equal to unity. The schematic representation of biomass is shown in Figure 4.69, where s the extracellular concentration of glucose which is both the carbon source, and S is the intracellular substrate concentration in pseudo-steady state equilibrium with s.

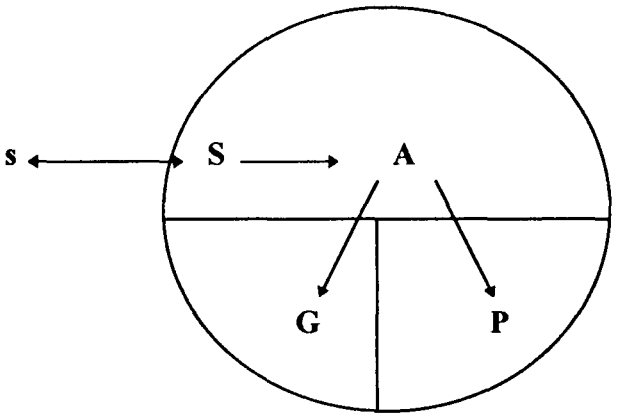


FIGURE 4.69. Structure of the three compartment model for fermentation of *E.coli* 294

The intracellular substrate S is used for the compartment A from which the formation of activated builing blocks of compartment A from which the two compartments (G and P) are synthesized. The reactions and their kinetics are:

$$S \longrightarrow \gamma_{11} X_A \qquad r_1 = \frac{k_1 S X_A}{S + K_1} \qquad (4.21)$$

$$X_A \longrightarrow \gamma_{22} X_G \qquad r_2 = \frac{k_2 S X_A}{S + K_2} \qquad (4.22)$$

$$X_A \longrightarrow \gamma_{33} X_P \qquad r_3 = \frac{k_3 S X_A}{S + K_3} \qquad (4.23)$$

The reactions are of first order with respect to the fraction of active compartment A, and saturation with respect to intracellular substrate concentration S is assumed. Under the fermentation conditions used for *E.coli* 294, it is reasonable to assume that only the number of plasmids equal to the normal copy number is formed in each cell cycle. Therefore, the rate of formation of the plasmid encoded product Xp can be expressed as in Equation (4.23).

Compartments G and P both consist of high molecular weight material. Nielsen et al. (1989, 1991a) have assumed that the saturation constants with respect to S for the

formation of the two compartments G and P from A are identical, i.e. $K_2 = K_3$. Similarly they have noted that the stoichiometric consants are also identical, i.e. $\gamma_{22} = \gamma_{33}$.

The mass balances for the three compartments are

$$dX/dt = \Gamma^T \mathbf{r} - \mu \mathbf{X} \tag{4.24}$$

where the specific growth rate μ is given by

$$\mu = (\mathbf{1})^T \Gamma^T \mathbf{r} = \gamma_{11} r_1 - (1 - \gamma_{22})(r_2 + r_3) \tag{4.25}$$

and the stoichiometric matrix is equal to

$$\Gamma^T = \begin{bmatrix} \gamma_{11} & -1 & -1 \\ 0 & \gamma_{22} & 0 \\ 0 & 0 & \gamma_{22} \end{bmatrix}$$

then the mass balances for the three compartments can be expressed as follows:

$$\frac{d}{dt}\begin{bmatrix} X_A \\ X_G \\ X_P \end{bmatrix} = \begin{bmatrix} \gamma_{11} & -1 & -1 \\ 0 & \gamma_{22} & 0 \\ 0 & 0 & \gamma_{22} \end{bmatrix} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} - \mu \cdot \begin{bmatrix} X_A \\ X_G \\ X_P \end{bmatrix} \tag{4.24a}$$

The mass balances for each copmpartment can be by using Equation (4.24a)

$$\frac{dX_A}{dt} = \gamma_{11} r_1 - r_2 - r_3 - \mu X_A \tag{4.26}$$

Inserting Equations (4.21-4.24) into Equation (4.26), gives:

$$\frac{dX_A}{dt} = \gamma_{11} k_1 X_A (1 - X_A)\left(\frac{S}{S + K_1}\right) - X_A (k_2 + k_3)(\gamma_{22} X_A - (1 - X_A))\left(\frac{S}{S + K_2}\right)$$

From Equation (4.24a),

$$\frac{dX_G}{dt} = \gamma_{22}\, r_2 - \mu\, X_G \qquad (4.27)$$

Inserting Equations (4.21-4.24) into Equation (4.27), gives:

$$\frac{dX_G}{dt} = X_A \left(\gamma_{22}\, k_2 - X_G \left(k_2 + k_3\right)\left(\gamma_{22} - 1\right)\right)\left(\frac{S}{S + K_2}\right) - \left[\frac{X_A\, X_G\, \gamma_{11}\, k_1\, S}{\left(S + K\right)}\right]$$

Finally, from Equation (4.24a)

$$\frac{dX_P}{dt} = \gamma_{22}\, r_3 - \mu\, X_P \qquad (4.28)$$

and inserting Equations (4.21-4.24) into Equation (4.28), gives:

$$\frac{dX_P}{dt} = \frac{-\gamma_{11}\, k_1\, X_A\, X_P S}{\left(S + K_1\right)} + X_A \left(\gamma_{22}\, k_3 - X_P \left(k_2 + k_3\right)\left(\gamma_{22} - 1\right)\right)\left(\frac{S}{S + K_2}\right)$$

The term γ_{22} specifies the yield coeffient for macromolecular synthesis from building blocks, and it is reported to be close to 1. The coefficient is smaller than 1 since compounds that are not included in the model are also formed by these reactions (e.g. water, acetic acid) (Nielsen et al., 1991a). A value of 0.95 is suggested for *E..coli* fermentation; however a sligthly smaller value, 0.9, was used in this work. The parameters, γ_{11} and k_1 were evaluated as a single parameter $\gamma_{11}\, k_1$ to decrease the number of parameters. The number of parameters to be estimated is reduced to 5 with these assumptions.

In order to convert the intracellular substrate to extracellular substrate, it is assumed that there is pseudo-steady state equilibrium between the intracellular substrate S and extracellular substrate s. The rate of uptake of substrate and the substrate concentration were assumed to be directly related to each other as in the two compartment model

proposed by Williams (Roels, 1983) (i.e., S= 0.90 s). It was assumed that 90 percent of the extracellular substrate can be utilized intracellularly. The inhibiton effect of substrate was not taken into account, since it was observed at high substrate concentrations in the experimental data obtained.

The kinetic parameters in the model, $\gamma_{11} k_1$, k_2 , k_3 , K_1 , K_2 , were evaluated by the nonlinear least squares method of Levenberg-Marquardt. The results were compared with the experimental data, recalculated in the optimization routine and fed again to the integration step until minimal error between experimental and integrated values was achieved. The computer programs which were used in the calculation of the three compartment model, are given in Appendix.... The parameters estimated from the model equations are as follows:

$\gamma_{11} k_1$	=	5.46978	(gA/ g hr)
k_2	=	10.5955	(1 / hr)
k_3	=	2.25786	(1 / hr)
K_1	=	2.68520	(g / L)
K_2	=	1.53891	(g / L)

Simulation results obtained for the variation mass fractions of compartment G and A are compared with the experiemntal data in Figures 4.70 and 4.71 respectively. The experimental X_A fraction was calculated from the measured RNA contents by considering that the RNA proportion of ribosomes is constant at 60%. The X_G fraction was calculated from the experiemntal data by substracting the measured RNA and main plasmid product contents from the total cell mass.

The three-compartment model including a number of simplifying assumptions gives a good general description of the growth kinetics with an initial glucose concentration of 10g/L. The simulations of Figures 4.70 and 4.71 indicate that the model is successful in predicting the fraction of the G compartment.The deviations observed in the fraction of the A compartment are most likely due to the over-simplified description of the substrate up-take as well as the fact that induction with IPTG has been overlooked. This deviation is reflected in the fraction of the plasmid product X_p, since $X_A + X_G + X_P = 1.0$.

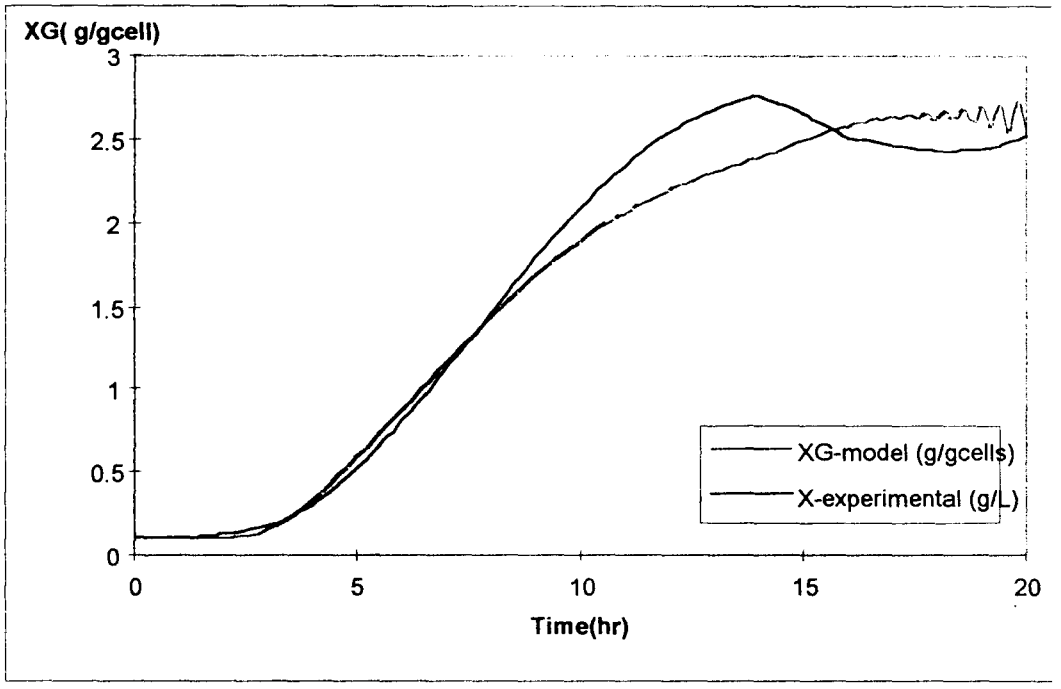


FIGURE 4.70. Comparison of experimental data and simulation of G-compartment (S_0= 10g/L)

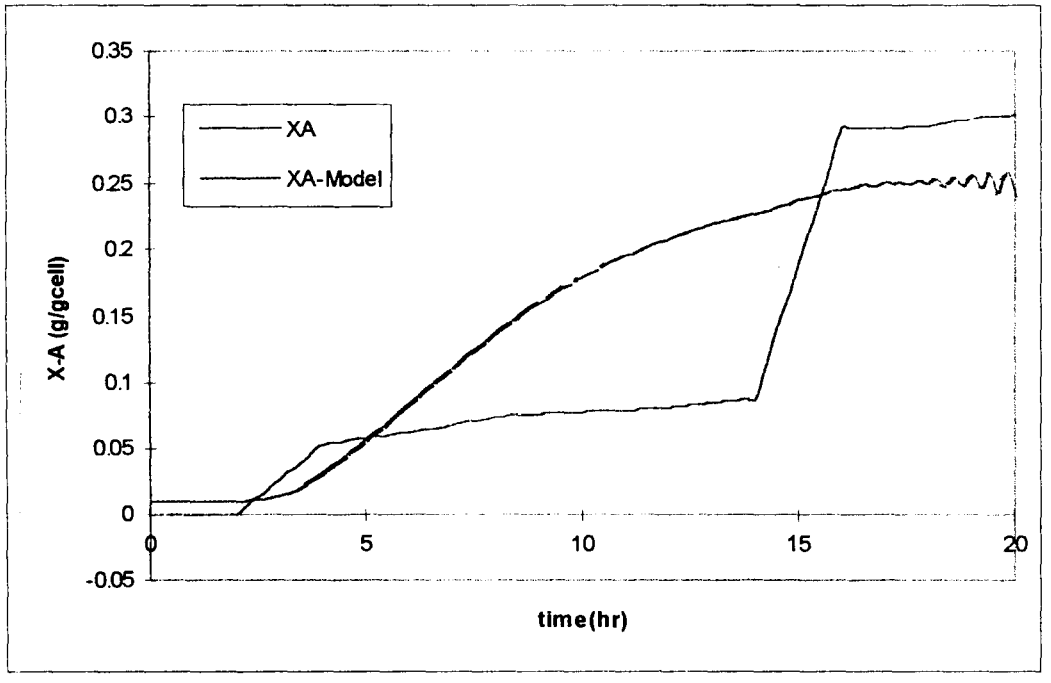


FIGURE 4.71. Comparison of experimental data (RNA/60%)and simulation of A compartment (S_0=10g/L)

The three-compartment model can be improved by changing the rate expression for substrate uptake; i.e either by using a relation of the Monod type or substrate-inhibited growth kinetics as given by Equation (4.17). Another improvement may be made by using two different rate expressions for the formation of X_A during the periods before and after induction with IPTG.

5. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The conclusions that can be drawn from the experimental findings of this study, can be summarized as follows:

(1) A laboratory scale procedure developed for the purification of *Eco*RI restriction endonuclease has been applied to two different *Escherichia* coli strains which are genetically modified to overproduce the enzyme. The purification method consist of three successive chromatographic steps including phosphocellulose and hydroxyapatite and followed by a further fractionation on a second phosphocellulose column. Quality control tests have indicated enzyme preparations free of contaminants and endo/exonucleases. The yields obtained at the final stage of the purification are $1.3x10^5$ U/g cells for *E.coli* M5248 and 3.3×10^6 U/gcells *E.coli* strain 294. The higher productivity of the latter strain is explained not only by the presence of a different promoter but also by the plasmid instability of the *E.coli* M5248 strain.

(2) Induction parameters such as inducer concentration, period of induction and the cell concentration at which inducer is to be added to the fermentation broth were optimized in order to increase the yield of the *Eco*RI restriction endonuclease isolated from recombinant *E.coli*. Plasmid pPG430 containing the *Eco*RI endonuclease and the methylase genes under the control of lac promoter was used in the experiments where induction was accomplished by using IPTG. 0.1 mM IPTG concentration at an optical density of 1.2 at 595 nm over an induction period of 6 hours were determined to be the optimum conditions for induction.

(3) For the recombinant *E.coli* M5248 (pSCC2) strain, the induction of *Eco*RI endonuclease was achieved by a temperature-upshift from 32°C to 42°C at an optical density of 1 at 590 nm. A slight decrease in the specific activity and the enzyme yield was observed for induction periods longer than 5 hours.

(4) The plasmid stability of pPG430 in the presence of selective and non-selective media was investigated in the case of *E.coli* 294. The effect of various culture conditions on the recombinant plasmid stability and consequently on the *Eco*RI productivity was determined. The maximum EcoRI activity was found when cells were cultured in L-broth with the addition of antibiotics. Minimal medium was found to result in a decrease in the enzyme productivity and in the plasmid stability.

(5) In the case of *E.coli* M5248 (pSCC2) cells, induction by temperature-upshift leads to a decrease in the percentage of plasmid-containing cells, which drops below 50% after 5 hours of exposure to elevated temperatures.

(6) Investigation of the stabilities of plasmid pPG430 and pSCC2 after successive generations showed that *E.coli* 294 cells containing plasmid pPG430 were more stable in maintaining their plasmids.

(7) The mathematical model used for comparing the specific growth rates of plasmid-free and plasmid-containing cells for both *E.coli* 294 and *E.coli* M5428 indicated that the plasmid stability is likely to be due to the growth rate differences rather than segregational instability.

(8) The experiments performed to investigate the effect of the change in substrate concentration on growth and product formation capability of the cells showed that both recombinant strains have similar growth behaviour with respect to different initial glucose concentrations under the same fermentation conditions. The inhibitory effect of glucose on growth when supplied at high concentrations (>15g/L) was observed in both recombinant *E.coli* M5248 and *E.coli* 294 cell cultures.

(9) The determination of the intracellular components of recombinant *E.coli* 294 cells containing plasmid pPG430, showed that the variation in RNA concentrations follows a trend similar to the growth curves of the cells. The plasmid DNA content of the cells were found to increase at the beginning of the exponential phase followed by a slight decrease in the amount of DNA when the concentration of RNA begins to increase. The total protein content were also found to increase following the induction of the recombinant cells.

(10) The plasmid copy number of average cells were determined by using a segregated model. The calculated copy numbers indicate that the increase in the copy numbers of both plasmids after induction.

(11) The unstructured modelling of cell growth has shown that substrate-inhibited kinetics is quite successful in predicting the growth data for both *E.coli* 294 and *E.coli* M5248. The model equation derived for *E.coli* 294 (pPG430) and *E.coli* M5248 (pSCC2) straina are as follows:

$$\mu = \frac{0.343 \cdot S}{0.832 + S} \cdot \left[1 - \frac{S}{30.2}\right]^{0.06}$$

$$\mu = \frac{0.371 \cdot S}{1.611 + S} \cdot \left[1 - \frac{S}{40.8}\right]^{0.24}$$

(12) A three compartment model was developed to describe the dynamic changes in the intracellular components of *E.coli* 294 cells. The model gives a good description of the growth kinetics with an initial glucose concentration of 10g /L. The model is successful in predicting the fraction of the G-compartment.

5.2 Recommendations

The purification protocols developed for the recovery of *Eco*RI endonuclease may be improved by replacing some of the chromatographic separation steps by rapid and effective HPLC applications.

New recombinant strains may be constructed in order to increase the productivity of *Eco*RI endonuclease or simplify its purification protocol.

Suitable model equations need to be selected for describing substrate utilization as a function of time and coupled with the substrate- inhibited biomass formation equations for simulating the x versus t and s versus t curves obtained from the fermentations conducted by using different initial glucose concntrations.

The three compartment model can be improved by changing the rate expression for substrate uptake; i.e either by using a relation of the Monod type or substrate-inhibited growth kinetics. Another improvement may be made in the three compartment model by using two different rate expressions for the formation of X_A during the periods before and after induction with IPTG.

It may be recommended for future studies that experiments may be performed in different reactor configurations such as fed-bacth and CSTR.

APPENDIX A. RELATION BETWEEN OPTICAL DENSITY AND DRY CELL WEIGHT

TABLE A.1. Dry cell weight data for *E.coli* 294 and *E.coli* M5248 cells

E.coli M5248		*E.coli* 294	
OD	**X-(g/L)**	**OD**	**X-(g/L)**
0.009	0.02	0.018	0.05
0.028	0.15	0.138	0.68
0.128	0.48	0.343	1.28
0.278	0.95	0.402	1.52
0.324	1.17	0.61	1.95
0.413	1.56	0.789	2.45
0.572	1.82	0.871	2.94
0.672	2.03	0.937	3.12
0.748	2.24	0.983	3.36
0.874	2.84	1.046	3.52
1.016	3.08	1.092	3.69
1.109	3.34	1.126	3.86
1.118	3.36	1.132	3.91

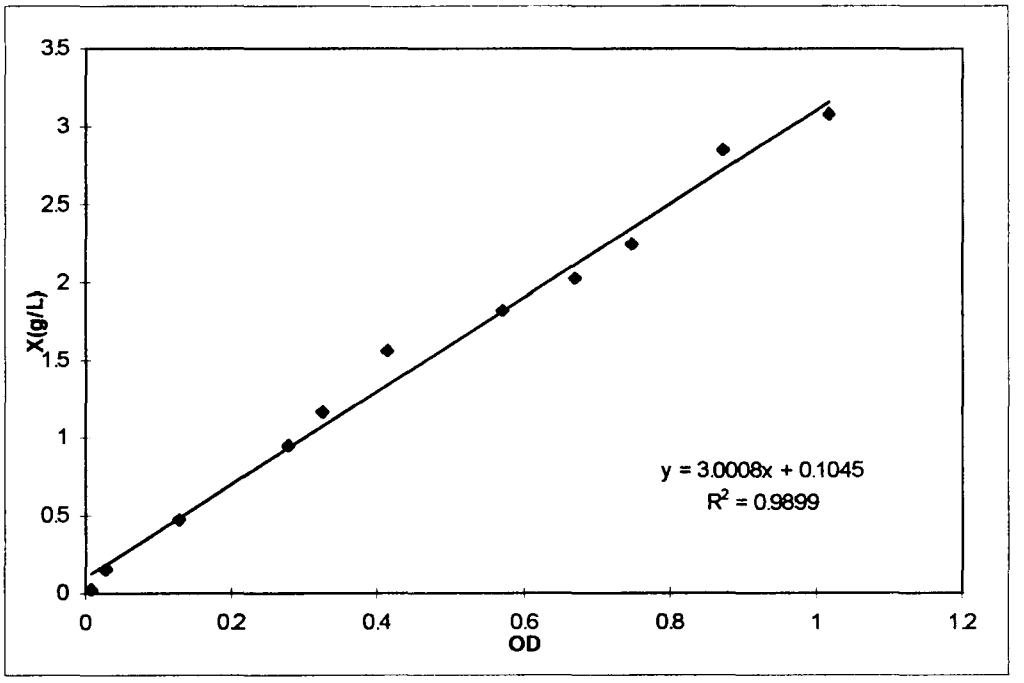


FIGURE A.1. Linear relation between dry cell weight and OD_{590} nm for *E.coli* M248 cells

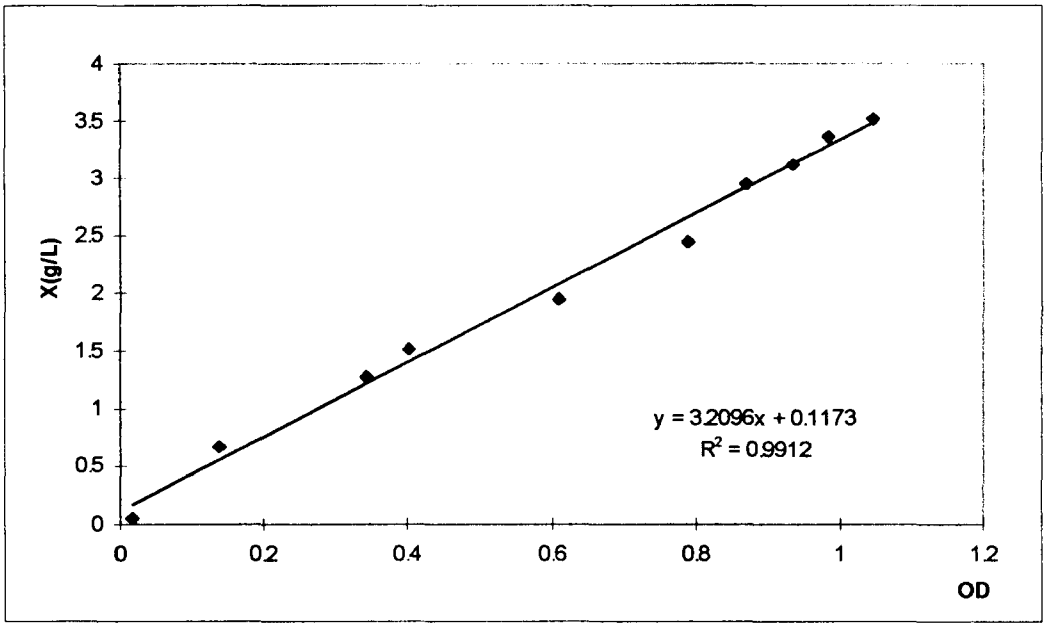


FIGURE A.2. Linear relation between dry cell weight and OD_{595} nm for *E.coli* 294 cells

APPENDIX B. CALIBRATION CHART FOR BRADFORD DYE BINDING METHOD

TABLE B.1. Standart data of Bradford assay

BSA (mg)	OD (595nm)	BSA (mg)	OD (595nm)
1	0.102	15	0.6
2	0.196	20	0.794
4	0.254	25	0.922
6	0.288	30	1.024
10	0.392		

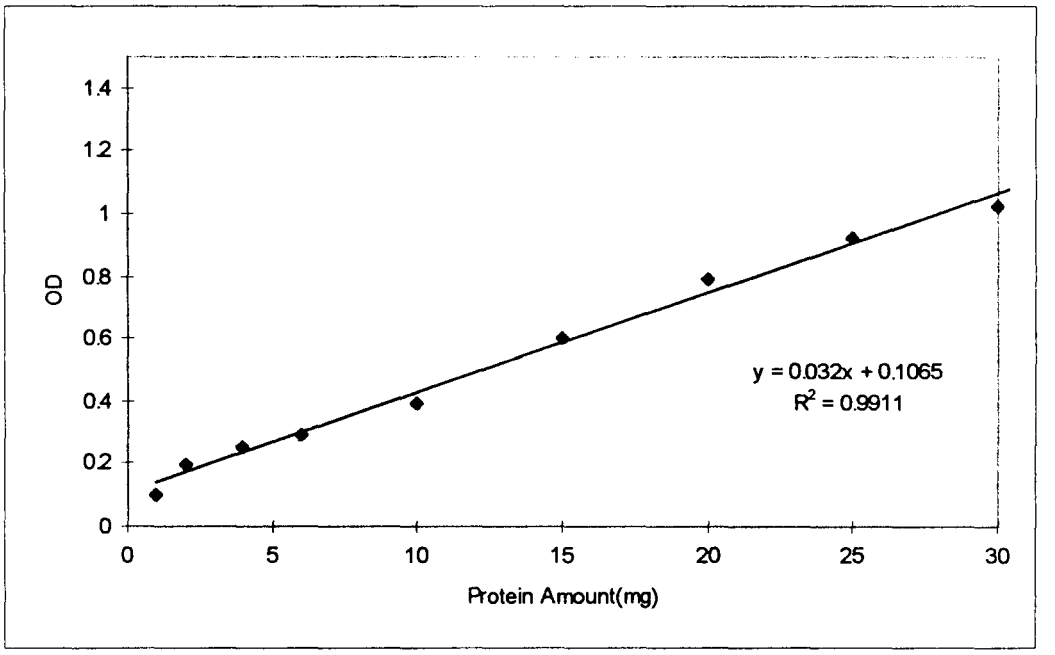


FIGURE B.1. Calibration chart for Bradford assay

APPENDIX C. EXPERIMENTAL DATA FOR BIOMASS AND SUBSTRATE CONCENTRATION

TABLE C.1. Experimental data for biomass and substrate concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations1, 5, 10 g/L.

S_0 = 1g/L			S_0 = 5g/L			S_0 = 10g/L		
time (hr)	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**
0	0.048	1	0	0.101	5.43	0	0.111	10
2	0.0844	0.92	2	0.157	5.01	2	0.124	9.77
4	0.278	0.72	4	0.312	3.73	4	0.374	7.55
6	0.426	0.54	7	0.844	2.54	6	0.828	6.33
8	0.702	0.39	11	2.106	1.02	8	1.571	5.04
10	1.008	0.26	13	2.445	0.64	10	2.31	4.32
12	1.172	0.19	16	2.787	0.44	12	2.724	3.5
14	1.366	0.14	18	2.994	0.26	14	2.803	2.24
16	1.399	0.05	20	3.033	0.19	16	2.882	1.72
18	1.405		21	3.05	0.11	18	2.954	0.85
						19	3.043	0.53
						20	3.079	0.36

TABLE C.2. Experimental data for biomass and substrate concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 15, 20, 30g/L.

S_0 = 15g/L			S_0 = 20g/L			S_0 = 30g/L		
time (hr)	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**
0	0.111	15	**0**	0.088	20.16	0	0.118	30
2	0.128	14.85	3	0.47	19.56	3	0.226	29.22
4	0.216	14.24	6	0.642	17.64	6	0.374	25.31
6	0.476	11.87	8	1.054	15.42	9	0.555	19.13
8	1.193	9.94	10	1.758	13.62	10	0.848	17.24
10	1.932	6.54	12	2.473	11.74	12	1.465	15.14
12	2.524	3.85	13	2.611	9.82	13	2.008	13.79
13	2.764	2.52	15	2.858	7.69	15	2.392	12.94
15	3.02	1.28	16	3.006	6.87	16	2.997	11.56
16	3.175	1.16	17	3.084	5.23	17	3.658	10.43
17	3.24	0.98	18	3.115	4.92	18	3.772	9.86
18	3.24	0.78	19	3.142	4.38	19	3.854	8.64
19	3.25	0.69	20	3.173	3.86	21	3.881	8.04
			21	3.197	3.37	22	3.907	7.64
			22	3.207	2.94			

TABLE C.3. Experimental data for biomass and substrate concentrations for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 1, 5, 10g/L.

S_0 = 1g/L			S_0 = 5g/L			S_0 = 10g/L		
time (hr)	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**
0	0.145	1.02	0	0.145	5.04	0	0.145	9.91
2	0.165	0.99	2	0.189	4.71	2	0.151	9.84
4	0.245	0.82	4	0.23	3.52	4	0.165	9.52
6	0.383	0.61	7	0.516	2.45	6	0.209	8.15
8	0.539	0.48	11	1.214	1.16	8	0.333	6.96
10	0.846	0.32	13	1.923	0.84	10	0.692	5.14
12	1.128	0.12	16	2.253	0.41	12	1.287	4.08
14	1.293	0.09	18	2.453	0.23	14	1.929	3.29
16	1.376	0.05	20	2.489	0.12	16	2.468	2.28
18	1.361		21	2.515	0.09	18	2.701	1.14
			23	2.548	0.05	19	2.748	0.94
						20	2.766	0.62
						21	2.777	0.41
						22	2.786	0.22

TABLE C.4. Experimental data for biomass and substrate concentrations for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentrations 15, 30 g/L.

S_0 = 15g/L			S_0 = 30g/L		
time (hr)	**X g/L**	**S g/L**	**time (hr)**	**X g/L**	**S g/L**
0	0.133	14.82	0	0.142	29.64
2	0.148	14.61	3	0.186	29.11
4	0.162	13.62	6	0.307	26.33
6	0.271	10.85	9	0.495	20.34
8	0.442	8.95	10	0.837	18.74
10	0.922	7.16	12	1.334	14.24
12	1.702	5.42	13	1.773	12.74
13	2.315	4.13	15	2.283	10.58
15	2.807	2.85	16	2.521	9.37
16	2.96	1.74	17	2.786	8.29
17	2.998	0.86	18	2.959	7.26
18	3.013	0.52	19	3.019	7.05
19	3.028	0.47	21	3.098	6.84
			22	3.107	6.62

APPENDIX D. EXPERIMENTAL DATA FOR CONCENTRATION OF INTRACELLULAR COMPONENTS

TABLE D.1. Experimental data for*Eco*RI activity, plasmid DNA, RNA and total protein concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 1 g/L.

S_0 = 1 g/L				
time(hour)	***Eco*RI Activity (U)**	**RNA Content (mg/ml)**	**DNA Content (mg/g cell)**	**Total Protein Amount (mg)**
0	0	0	0	0
2	0	0	0.108	0
4	0	0.007	0.165	2.575
6	0	0.013	0.142	8.37
8	0	0.025	0.129	14.026
10	23345	0.039	0.121	25.641
12	100050	0.052	0.118	42.354
14	300150	0.211	0.116	196.248
16	533600	0.228	0.115	229.248
18	833750	0.234	0.114	242.836

TABLE D.2. Experimental data for*Eco*RI activity, plasmid DNA, RNA and total protein concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 5 g/L.

S_0 = 5 g/L				
time(hour)	***Eco*RI Activity (U)**	**RNA Content (mg/ml)**	**DNA Content (mg/g cell)**	**Total Protein Amount (mg)**
0	0	0	0	0
2	0	0	0.152	0
4	0	0.01	0.236	4.926
7	0	0.031	0.217	18.431
11	53360	0.081	0.203	62.543
13	83375	0.107	0.199	75.238
16	300150	0.389	0.196	336.452
18	700350	0.496	0.195	369.781
20	1500750	0.518	0.194	421.567
21	1834250	0.527	0.193	458.654

TABLE D.3. Experimental data for*Eco*RI activity, plasmid DNA, RNA and total protein concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 10 g/L.

S_0 = 10 g/L				
time(hour)	***Eco*RI Activity (U)**	**RNA Content (mg/ml)**	**DNA Content (mg/g cell)**	**Total Protein Amount (mg)**
0	0	0	0	0
2	0	0	0.157	0
4	0	0.012	0.267	5.364
6	0	0.031	0.259	19.327
8	30015	0.07	0.25	57.849
10	50025	0.108	0.245	70.676
12	66700	0.131	0.232	78.931
14	100050	0.148	0.229	84.527
16	833750	0.502	0.224	386.118
18	1667500	0.518	0.221	477.391
19	2001000	0.546	0.219	497.644
20	2134400	0.556	0.218	516.879

TABLE D.5. Experimental data for*Eco*RI activity, plasmid DNA, RNA and total protein concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 20 g/L.

S_0 = 20 g/L				
time(hour)	***Eco*RI Activity (U)**	**RNA Content (mg/ml)**	**DNA Content (mg/g cell)**	**Total Protein Amount (mg)**
0	0	0	0	0
3	0	0.014	0.264	6.692
6	0	0.022	0.377	18.244
8	0	0.041	0.341	34.651
10	300150	0.072	0.319	57.389
12	600300	0.105	0.307	73.633
13	833750	0.118	0.295	81.576
15	1500750	0.134	0.287	90.964
16	3001500	0.152	0.281	98.362
17	6003000	0.522	0.272	408.313
18	1.7E+07	0.583	0.268	458.517
19	2.7E+07	0.638	0.26	479.129
20	3.3E+07	0.684	0.255	492.524
21	3.8E+07	0.692	0.253	497.628
22	4.4E+07	0.705	0.249	500.34

TABLE D.6. Experimental data for*Eco*RI activity, plasmid DNA, RNA and total protein concentrations for *E.coli* 294 (pPG430) cells grown in minimal medium with initial glucose concentrations 30 g/L.

S_0 = 30 g/L				
Time(hour)	***Eco*RI Activity (U)**	**RNA Content (mg/ml)**	**DNA Content (mg/g cell)**	**Total Protein Amount (mg)**
0	0	0	0	0
3	0	0.003	0.208	0
6	0	0.012	0.282	5.125
9	0	0.024	0.348	13.575
10	0	0.041	0.331	24.65
12	26680	0.074	0.32	52.85
13	50025	0.104	0.314	62.4
15	75037.5	0.142	0.308	71.25
16	166750	0.186	0.305	81.175
17	500250	0.622	0.302	367.15
18	1667500	0.733	0.3	393.125
19	3001500	0.779	0.299	420.975
21	3335000	0.784	0.299	440
22	3601800	0.796	0.298	445.9
22		0.705	0.249	500.34

TABLE D.7. Experimental data for*Eco*RI activity for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentration of 1, 5, 10 g/L.

S_0 =	1g/L	S_0 =	5g/L	S_0 =	10g/L
Time(hour)	***Eco*RI Activity (U)**	**Time(hour)**	***Eco*RI Activity (U)**	**Time(hour)**	***Eco*RI Activity (U)**
0	0	0	0	0	0
2	0	2	0	2	0
4	0	4	0	4	0
6	0	7	0	6	0
8	0	11	0	8	0
10	0	13	0	10	0
12	0	16	0	12	0
14	66700	18	66700	14	0
16	300150	20	200100	16	0
18	500250	21	600300	18	80040
		23	1167250	19	300150
				20	500250
				21	1000500
				22	1334000

TABLE D.8. Experimental data for*Eco*RI activity for *E.coli* M5248 (pSCC2) cells grown in minimal medium with initial glucose concentration of 15, 30 g/L.

S_0 =	**15g/L**	S_0 =	**30g/L**
Time(hour)	***Eco*RI Activity (U)**	**Time(hour)**	***Eco*RI Activity (U)**
0	0	0	0
2	0	3	0
4	0	6	0
6	0	9	0
8	0	10	0
10	0	12	0
12	0	13	0
13	66700	15	0
15	200100	16	0
16	500250	17	66700
17	1000500	18	133400
18	1334000	19	500250
19	1667500	21	667000
		22	1000500

APPENDIX E. CURVE FITTING TECHNIQUE AND COMPUTER PROGRAM

The number of experimental data points that can be taken at regular time intervals during an experiment is rather limited. In modelling, however, it is important to use a large number of data points with homogenous time intervals. Arbitrary mathematical functions flexible enough to fit the set of experimental data on biomass formation (x versus t) and substrate consumption (s consumed versus t) for both recombinant *E.coli* strains and intracellular components in *E.coli* 294 were used for generating new data points by interpolation so that data at 0.2 hr time intervals could be used in model discrimination. In the case of x versus t data, this technique introduced some error and therefore was not used in unstructured modelling. The curve fitting program applied to data obtained at different initial glucose concentrations and the curves for x vs t, s vs t and RNA vs t are given below.

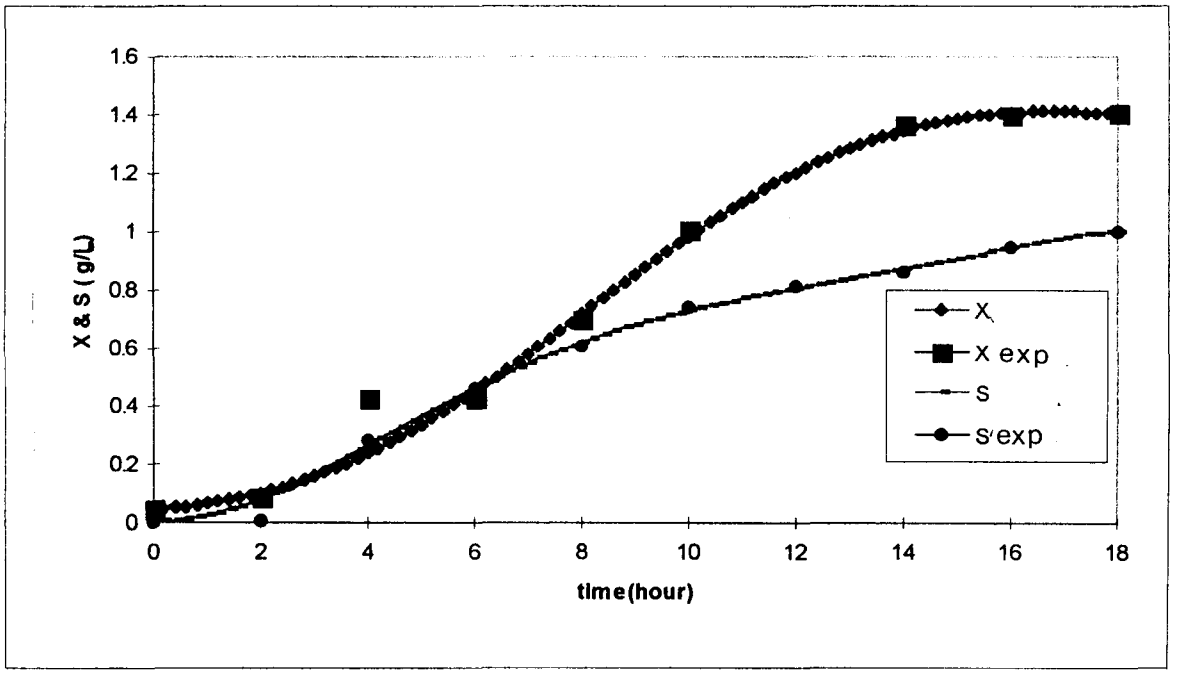


FIGURE E.1. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 1g/L)

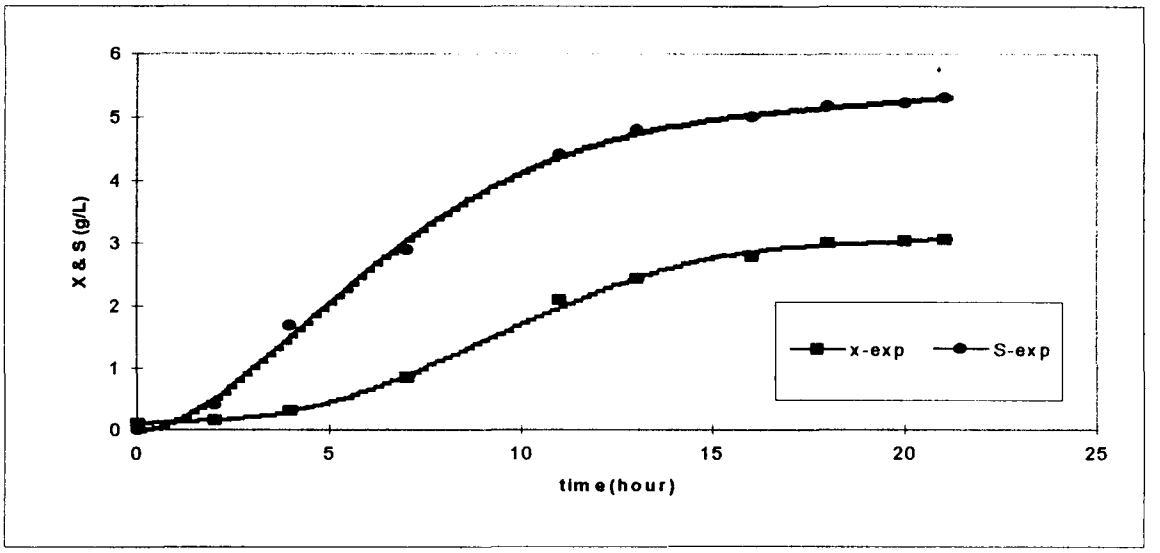


FIGURE E.2. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 5g/L)

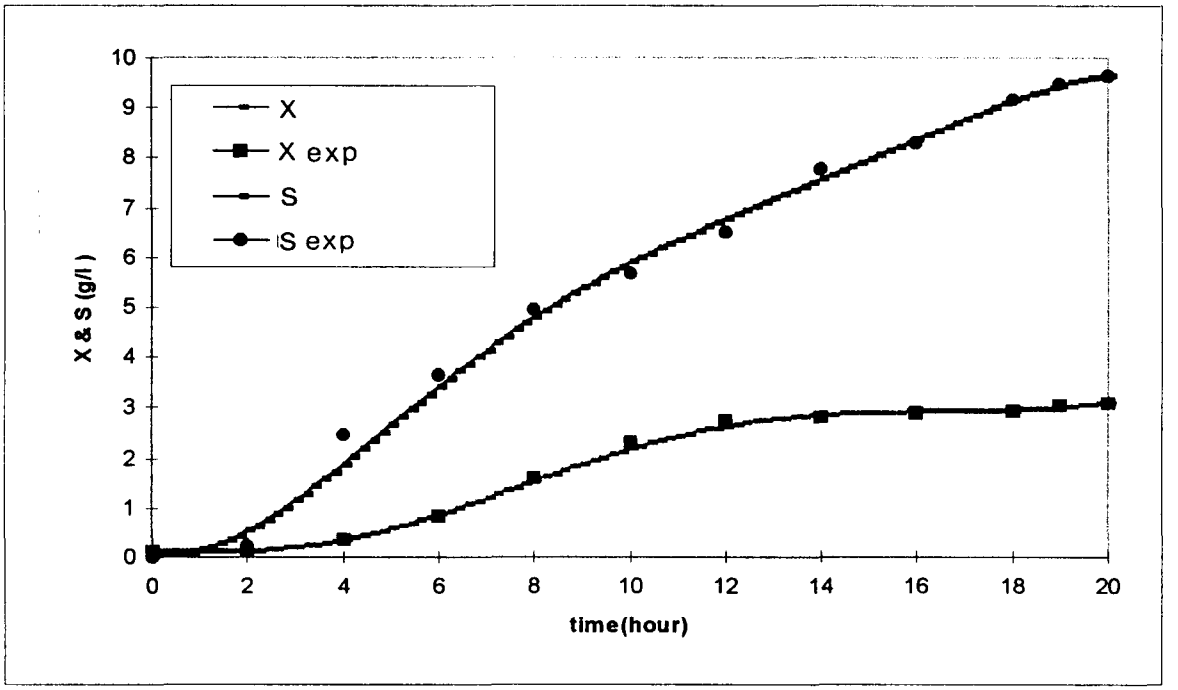


FIGURE E.3. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 10g/L)

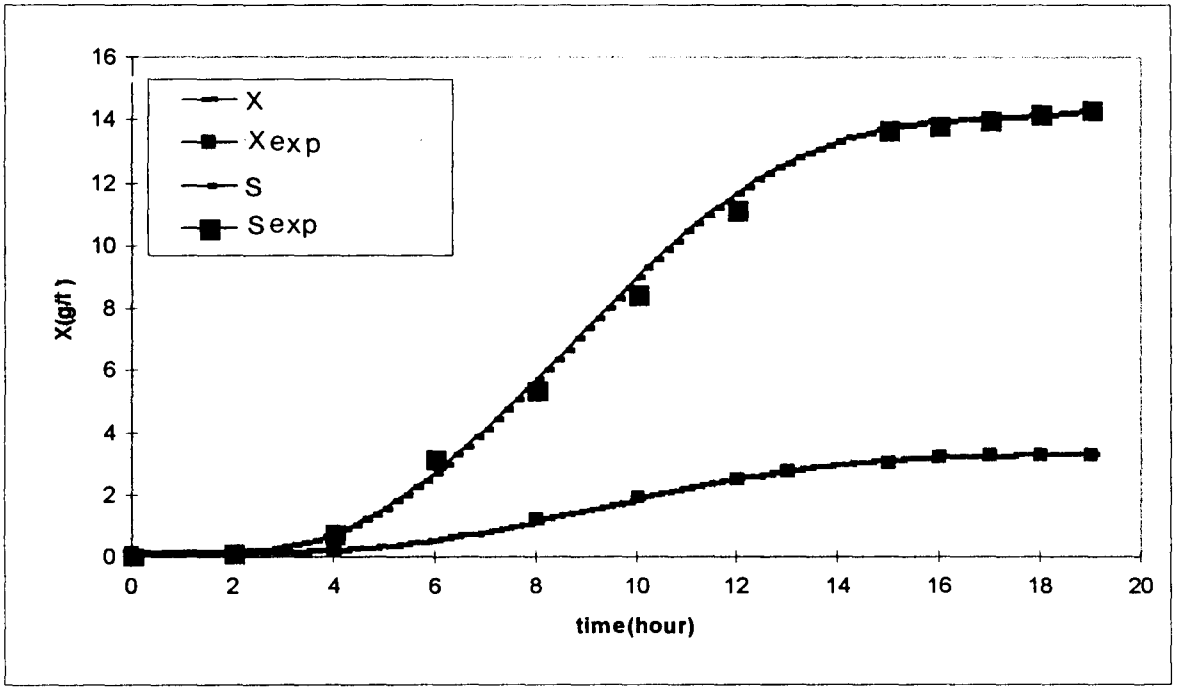


FIGURE E.4. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 15g/L)

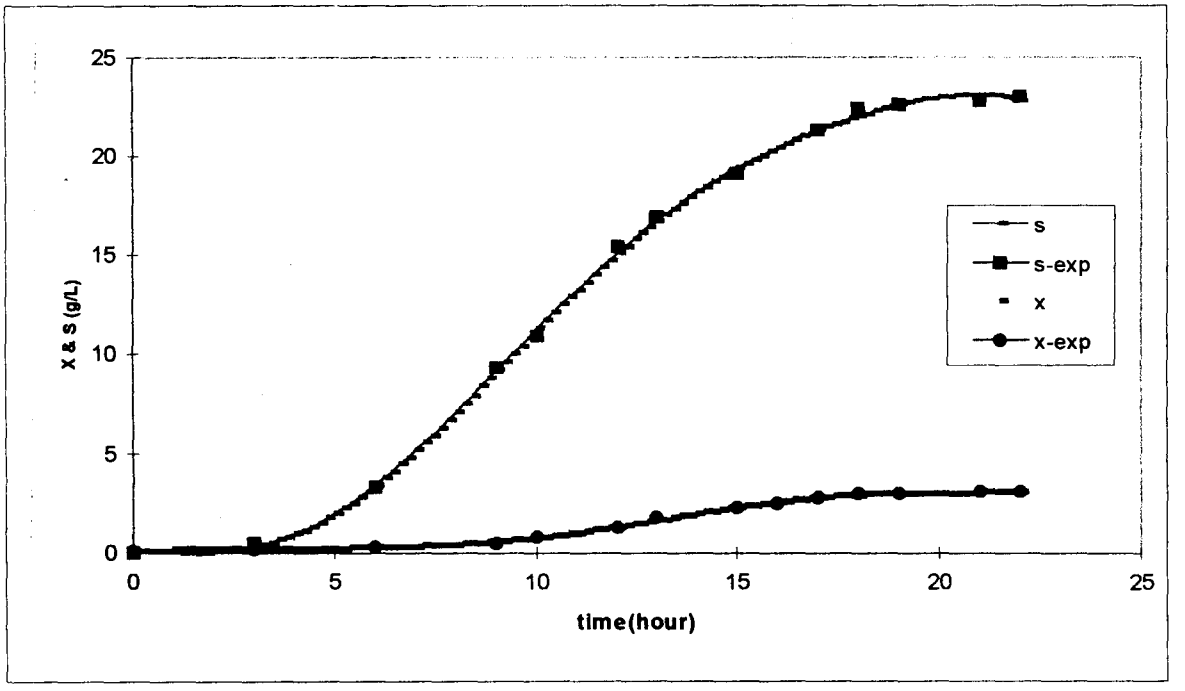


FIGURE E.5. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* 294 (S_0 = 30g/L)

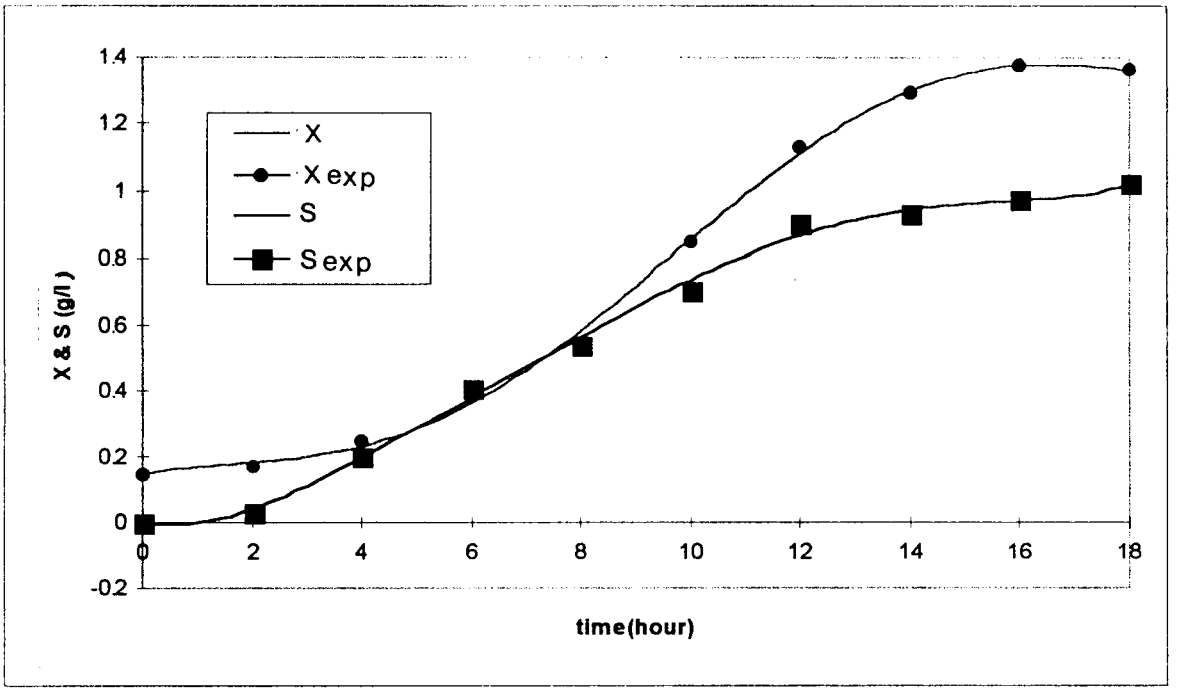


FIGURE E.6. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 1g/L)

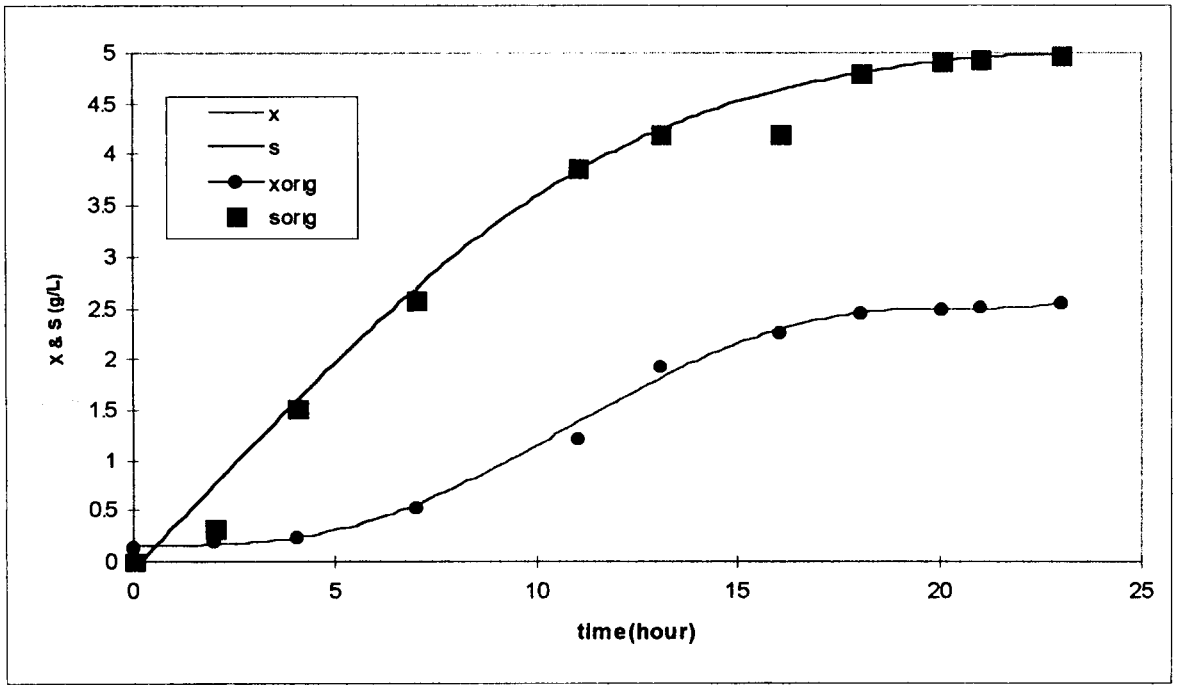


FIGURE E.7. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 5g/L)

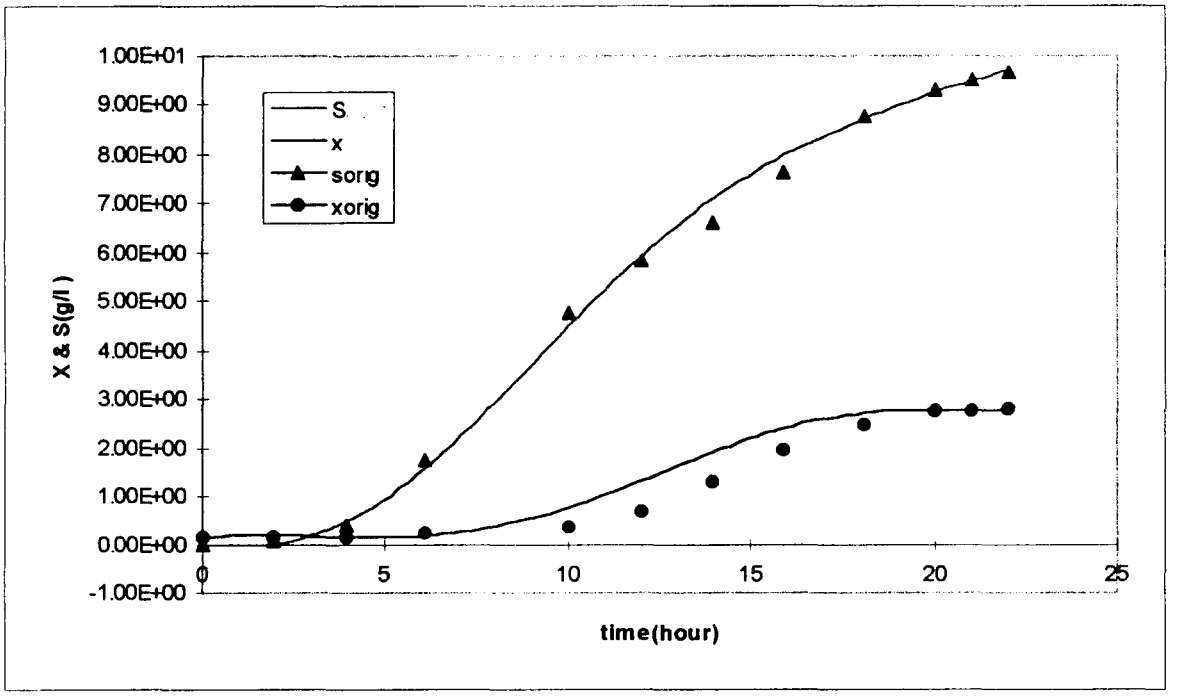


FIGURE E.8. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 10g/L)

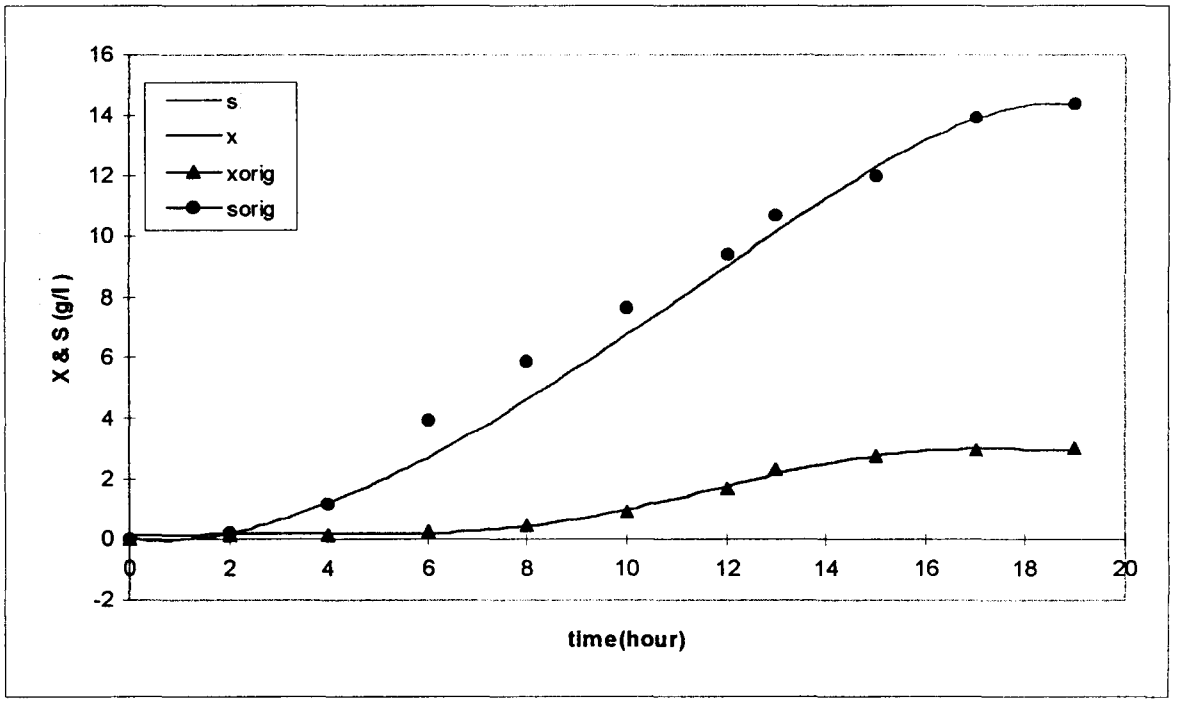


FIGURE E.9. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248(S_0 = 15g/L)

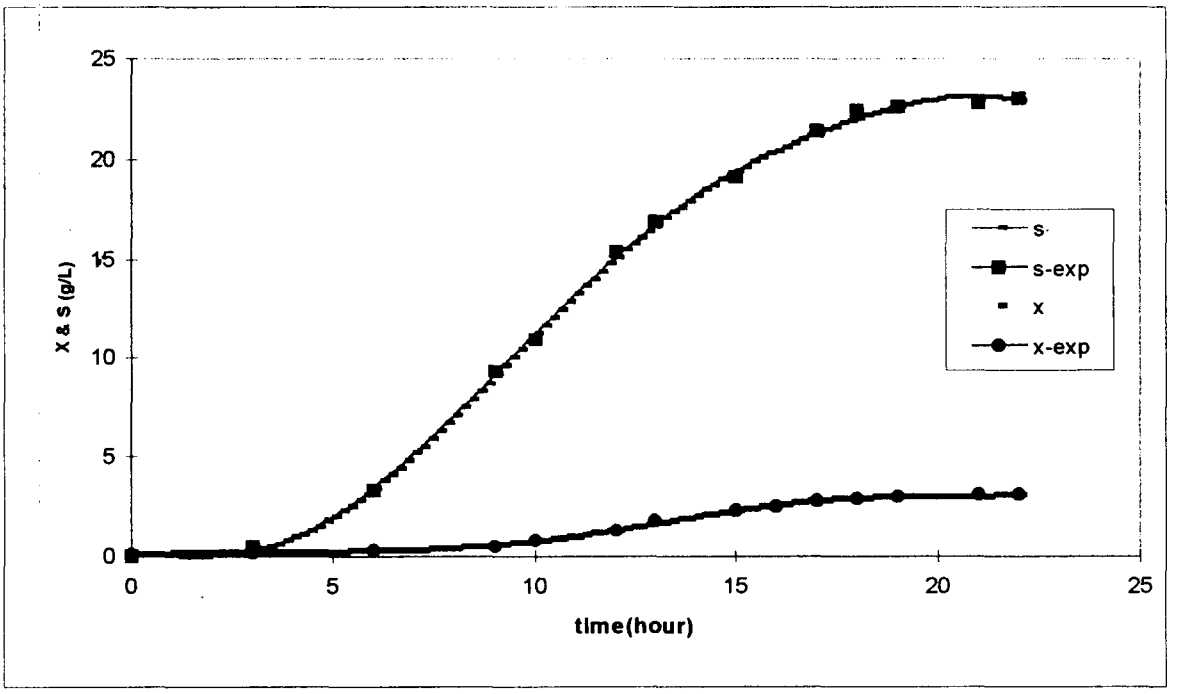


FIGURE E.10. Curve fitting applied to the biomass formation and the glucose consumption data for *E.coli* M5248 (S_0 = 30g/L)

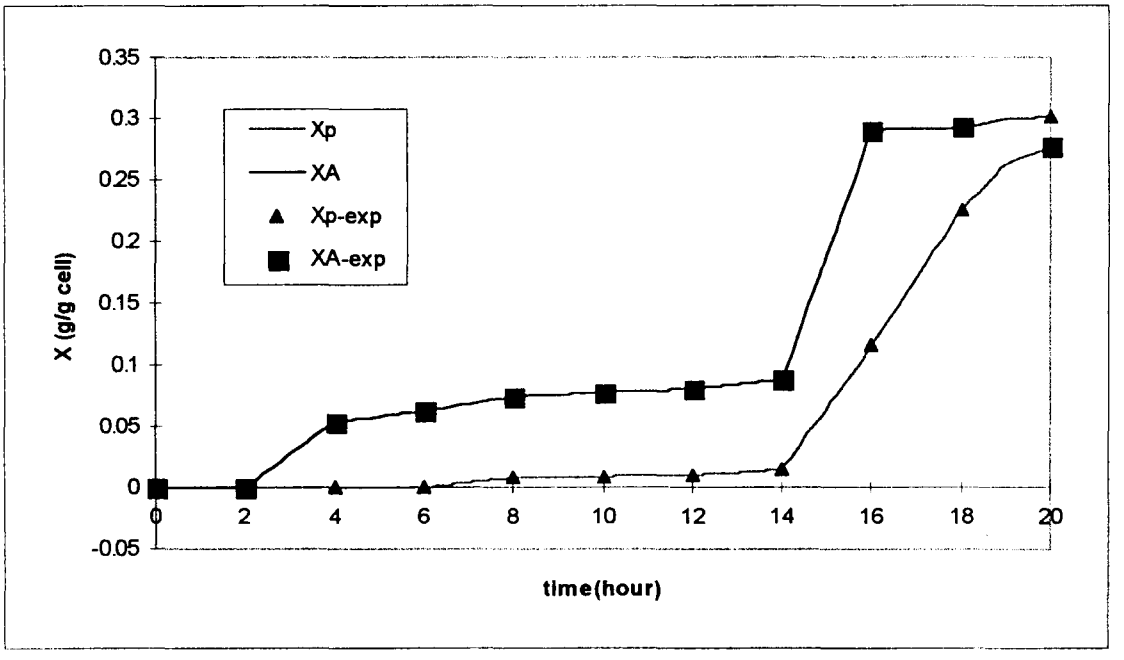


FIGURE E.11. Curve fitting applied to RNA and EcoRI formation data for *E.coli* 294 (S_0 = 10g/L)

In programming, numerical recipies subrotuines were used as sources (Press et al., 1988).

```
/* curve fitting program */
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <nr.h>
#include <nrutil.h>
#include <nrutil.c>
#include <gaussj.c>
#include <covsrt.c>
#include <fpoly.c>
#include <fleg.c>
#include <lfit.c>
main(argc,argv)
int argc;
char *argv[];
{
      int i,j,isum,nsteps,*count,binary,i_fit_func,order,iswap,i_imp,i_order;
        int ndata,ma,mfit,*lista;
        float *x,*y,*sig,*a,*p,**covar,chisq;
      char datain[20],dataout[40];
      float *ycoord, *xcoord, min, max, stepsize,base;
      FILE *in,*out;
        void (*fit_func)();
      if(argc!=9){
            printf("\nenter:  flnm datain order nsteps i_fit_func iswap importance base
change_order");
            printf("\ndatain is the name of the input data file");
            printf("\norder is the order of the fitting polynomial");
            printf("\nnpoints is # of smooth data points");
            printf("\nenter i_fit_func [1] : for polynomial fitting");
```

```
        printf("\n              [2] : for Legendre function fitting");
        printf("\niswap is 1 to change value and category axes else 0");
        printf("\nenter importance [1] : for high values");
        printf("\n              [2] : for difference from center");
        printf("\n              [12] : for difference from center (end point 0 variance)");
        printf("\n              [22] : for difference from center (first point 0 variance)");
        printf("\nbase is variance power base");
        printf("\nchange_order is  [1] to change order of s (else [0])");
printf("\n");
        exit(1);
  }
    ndata=30;
  strcpy(datain,argv[1]);
    ma = atoi(argv[2]);
    nsteps = atoi(argv[3]);
    i_fit_func = atoi(argv[4]);
    iswap = atoi(argv[5]);
    i_imp = atoi(argv[6]);
    base = atof(argv[7]);
    i_order = atoi(argv[8]);
    mfit = ma ;
    if(i_fit_func==1) fit_func=fpoly;
    else if(i_fit_func==2) fit_func=fleg;
    else {printf("\nenter a proper number for fitting polynomial (1,2)\n"); exit(1);}
    x = vector(1,ndata);
    y = vector(1,ndata);
    sig = vector(1,ndata);
    a = vector(1,ma);
    p = vector(1,ma);
    lista = ivector(1,ma);
    covar = matrix(1,ma,1,ma);
    xcoord = vector(0,nsteps);
    ycoord = vector(0,nsteps);
```

```
        i=0;
        if((in=fopen(datain,"rt"))==NULL)exit(1);
        while(!feof(in)){
                i++;
                if(iswap==1) fscanf(in,"%g %g",&y[i],&x[i]);
                else fscanf(in,"%g %g",&x[i],&y[i]);
          }
        fclose(in);
        ndata=i-1;
          printf("\ni=%d\tx=%lg",ndata,x[ndata]);
if(i_order==1){
  for(i=ndata; i>=1; i--)y[i] = y[1] - y[i] ;
  sprintf(dataout,"%s_reflected",argv[1]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",x[i],y[i]);
  fclose(out);
}
        min=5.e20;
          for(i=1; i<=ndata; i++) if(x[i]<min) min=x[i];
          max=-5.e20;
        for(i=1; i<=ndata; i++) if(x[i]>max) max=x[i];
          stepsize = (max-min)/ (float)nsteps ;
          xcoord[0] = min ;
          for(i=1; i<=nsteps; i++){
                xcoord[i] = xcoord[i-1] + stepsize ;
          }
          for(i=1; i<=ndata; i++) {
if((i_imp % 10) ==1)
                sig[i] = pow(base,y[i]);
else if((i_imp % 10) ==2)
                sig[i] = pow(base,fabs(y[i]-y[ndata/2]));
                sig[i] = 1. / sig[i] ;
                printf("\nsig[%d] = %f ",i,sig[i]);
```

```
        }
if(i_imp==12) sig[ndata]=0.0001;
if(i_imp==22) sig[1] = 0.0001;
        printf("\n");
        for(i=1; i<=ma; i++) lista[i] = i;
        lfit(x,y,sig,ndata,a,ma,lista,mfit,covar,&chisq,fit_func);
        for(i=1; i<=ma; i++) printf("\na[%d] = %f ",i,a[i]);
        printf("\n");
        for(i=0; i<=nsteps; i++){
                ycoord[i] = 0.;
                fit_func(xcoord[i],p,ma);
                for(j=1; j<=ma; j++) ycoord[i] += p[j] * a[j] ;
        }
        sprintf(dataout,"%s_o%s_n%s_f%s_sw%s_vi%s_vb%s",argv[1],argv[2],argv[3],arg
v[4],argv[5],argv[6],argv[7]);
      if((out=fopen(dataout,"wt"))==NULL)exit(1);
        if(iswap==1) for(i=0; i<=nsteps; i++) fprintf(out,"%g\t%g\n",ycoord[i],xcoord[i]);
        else for(i=0; i<=nsteps; i++) fprintf(out,"%g\t%g\n",xcoord[i],ycoord[i]);
        fclose(out);
}
```

APPENDIX F. COMPUTER PROGRAMS USED IN UNSTRUCTURED MODELING

In programming, subroutines of numerical recipies were used as source files (Press et al.,1988)

```
/* Unstructured Modeling Program */
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <nr.h>
#include <nrutil.h>
float *x,*s,*t;
void fcandan1();
void fcandan2();
void fcandan3();
void fcandan4();
void xfun_1();
void xfun_2();
void xfun_3();
void xfun_4();
#include <nrutil.c>
#include <gaussj.c>
#include <covsrt.c>
#include <mrqmin.c>
#include <mrqcof.c>
main(argc,argv)
int argc;
```

```
char *argv[];
{
 int i,j,isum,nsteps,*count,binary,i_fit_func,order,iswap=0,i_imp,i_prec,Sm;
 int ndata,ma,mfit,*lista;
 float *sig,*a,*p,**covar,**alpha,chisq,alamda,yfit,*dyda,*a_last,*a_init;
 float dummy;
 char data_x[80],data_s[80],dataout[100],initdata[20];
 float *ycoord, *xcoord, min, max, stepsize,base;
 int idir=1;
 float *x_new;
 FILE *in,*out;
 void (*fit_func)();
 void (*x_func)();
 if(argc!=6){
   printf("\nenter:  flnm data_x data_s i_fit_func nsteps parameter_file");
printf("\n");
   exit(1);
 }
 i_prec=3;
 ndata=300;
 strcpy(data_x,argv[1]);
 strcpy(data_s,argv[2]);
 i_fit_func = atoi(argv[3]);
 if(i_fit_func==1) ma = 4;
 else if(i_fit_func==2) ma = 3;
 else if(i_fit_func==3) ma = 4;
 else if(i_fit_func==4) ma = 3;
 nsteps = atoi(argv[4]);
x = vector(0,ndata+1);/*  !!!!!!! dikkat !!!!!!!!!!!!!!! */
 x_new = vector(0,ndata+1);
t = vector(0,ndata+1);
 s = vector(1,ndata);
 dyda = vector(1,ndata);
```

```
  sig = vector(1,ndata);
  a = vector(1,ma);
  a_last = vector(1,ma);
  a_init = vector(1,ma);
  p = vector(1,ma);
  lista = ivector(1,ma);
  covar = matrix(1,ma,1,ma);
  alpha = matrix(1,ma,1,ma);
  xcoord = vector(0,nsteps);
  ycoord = vector(0,nsteps);
if(i_fit_func==1){
   fit_func=fcandan1;  x_func = xfun_1 ;sprintf(initdata,"initf1");
  }
  else if(i_fit_func==2){
   fit_func=fcandan2;  x_func = xfun_2 ;sprintf(initdata,"initf2");
  }
  else if(i_fit_func==3){
   fit_func=fcandan3;  x_func = xfun_3 ;sprintf(initdata,"initf3");
  }
  else if(i_fit_func==4){
   fit_func=fcandan4;  x_func = xfun_4 ;sprintf(initdata,"initf4");
  }
  else {printf("\nenter a proper number for fitting polynomial (1,2)\n"); exit(1);}
i=0;
  if((in=fopen(data_x,"rt"))==NULL)exit(1);
  while(!feof(in)){
   i++;
   fscanf(in,"%g %g",&t[i],&x[i]);
  }
  fclose(in);
  ndata=i-1;
  t[0] = (2*t[1]) - t[2] ; t[ndata+1] = (2*t[ndata]) - t[ndata-1];
  x[0] = (2*x[1]) - x[2] ; x[ndata+1] = (2*x[ndata]) - x[ndata-1];
```

```
  if((in=fopen(data_s,"rt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++){
    fscanf(in,"%g %g",&dummy,&s[i]);
    if(dummy!=t[i]) {printf("\ntime data are different\n"); exit(1);}
  }
  fclose(in);
strcpy(initdata,argv[5]);
    if((in=fopen(initdata,"rt"))==NULL)exit(1);
    for(i=1; i<=ma; i++){
      fscanf(in,"\ta[%d]=%g",&i,&a_last[i]);
      printf("\ta[%d]=%g",i,a_last[i]);
    }
    fclose(in);
  x_new[0] = x[0]; x_new[1] = x[1] ;
  idir = 1;
  for(i=1; i<=ndata; i++) x_func(i,a_last,ma,idir,x_new) ;
sprintf(dataout,"%s..%s..%s..d%d",argv[1],argv[2],argv[5],idir);
        if((out=fopen(dataout,"wt"))==NULL)exit(1);
          if(iswap==1) for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
          else for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
          fclose(out);
  x_new[ndata] = x[ndata]; x_new[ndata+1] = x[ndata+1] ;
  idir = 0;
  for(i=ndata; i>=1; i--) x_func(i,a_last,ma,idir,x_new) ;
  sprintf(dataout,"%s..%s..%s..d%d",argv[1],argv[2],argv[5],idir);
        if((out=fopen(dataout,"wt"))==NULL)exit(1);
          if(iswap==1) for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
          else for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
          fclose(out);
}
void xfun_1(i,a,na,idir,x)
float a[],x[];
int i,na,idir;
```

```
{
  float frac,brac,power,mu;
  frac =  (s[i] * a[1]) / (a[3] + s[i]);
  brac = (1 - (s[i] / a[2]));
  power = pow(brac , a[4]);
  mu = frac * power;
  if(idir==1) x[i+1] = ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i-1];
  else if(idir==0) x[i-1] = - ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i+1];
}
void fcandan1(i,a,xmod,dxda,na)
float a[],*xmod,dxda[];
int i,na;
{
  float frac,brac,power,mu,factor; int j;
  frac = (s[i] * a[1]) / (a[3] + s[i]);
  brac = (1 - (s[i] / a[2]));
  power = pow(brac , a[4]);
  mu = frac * power;
  *xmod = (x[i+1] - x[i-1]) / (t[i+1] - t[i-1]) / mu ;
  factor = -(*xmod) / mu ;
  dxda[1] = mu / a[1] ;
  dxda[2] = frac * a[4] * pow(brac , (a[4] - 1)) * s[i] / a[2] / a[2] ;
  dxda[3] = mu * (-1. / (a[3] + s[i])) ;
  dxda[4] = mu * log(brac) ;
  for(j=1; j<=na; j++) dxda[j] *= factor ;
}
void xfun_2(i,a,na,idir,x)
float a[],x[];
int i,na,idir;
{
  float e1,e2,e3,mu;
  e1 = exp( -s[i] / a[2] );
  e2 = - exp( -s[i] / a[3] );
```

```
  e3 = e1 + e2 ;
    mu = a[1] * e3 ;
  if(idir==1) x[i+1] = ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i-1];
  else if(idir==0) x[i-1] = - ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i+1];
}
void fcandan2(i,a,xmod,dxda,na)
float a[],*xmod,dxda[];
int i,na;
{
  float e1,e2,e3,mu,factor; int j;
  e1 = exp( -s[i] / a[2] );
  e2 = - exp( -s[i] / a[3] );
  e3 = e1 + e2 ;
    mu = a[1] * e3 ;
  *xmod = (x[i+1] - x[i-1]) / (t[i+1] - t[i-1]) / mu ;
  factor = -(*xmod) / mu ;
  dxda[1] = e3 ;
  dxda[2] = a[1] * e1 * s[i] / a[2] / a[2] ;
  dxda[3] = a[1] * e2 * s[i] / a[3] / a[3] ;
  for(j=1; j<=na; j++) dxda[j] *= factor ;
}
void xfun_3(i,a,na,idir,x)
float a[],x[];
int i,na,idir;
{
  float N , brac1 , brac2 , D , mu ;
  N = a[1] * s[i] ;
  brac1 = s[i] * s[i] / a[2] ;
  brac2 = 1 + (s[i] / a[4]) ;
  D = a[3] + s[i] + (brac1 * brac2) ;
  mu = N/D ;
  if(idir==1) x[i+1] = ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i-1];
  else if(idir==0) x[i-1] = - ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i+1];
```

```
}
void fcandan3(i,a,xmod,dxda,na)
float a[],*xmod,dxda[];
int i,na;
{
 float N , brac1 , brac2 , D , mu , factor; int j;
 N = a[1] * s[i] ;
 brac1 = s[i] * s[i] / a[2] ;
 brac2 = 1 + (s[i] / a[4]) ;
 D = a[3] + s[i] + (brac1 * brac2) ;
 mu = N/D ;
 *xmod = (x[i+1] - x[i-1]) / (t[i+1] - t[i-1]) / mu ;
 factor = -(*xmod) / mu ;
 dxda[1] = s[i] / D ;
 dxda[3] = mu  / (-D);
 dxda[2] = dxda[3] * brac2 * brac1 / (-a[2]) ;
 dxda[4] = dxda[3] * brac1 * ( -s[i]) / a[4] / a[4] ;
 for(j=1; j<=na; j++) dxda[j] *= factor ;
}
void xfun_4(i,a,na,idir,x)
float a[],x[];
int i,na,idir;
{
 float e1,frac, mu ;
 e1 = exp (-s[i] / a[2]);
 frac = (a[1] * s[i]) / (a[3] + s[i]) ;
 mu = frac * e1 ;
 if(idir==1) x[i+1] = ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i-1];
 else if(idir==0) x[i-1] = - ((x[i] * mu) * (t[i+1]-t[i-1])) + x[i+1];
}
void fcandan4(i,a,xmod,dxda,na)
float a[],*xmod,dxda[];
int i,na;
```

```
{
 float e1,frac, mu , factor; int j;
 e1 = exp (-s[i] / a[2]);
 frac = s[i] / (a[3] + s[i]) ;
 dxda[1] = frac * e1 ;
 frac *= a[1] ;
 mu = frac * e1 ;
 *xmod = (x[i+1] - x[i-1]) / (t[i+1] - t[i-1]) / mu ;
 factor = -(*xmod) / mu ;
   dxda[2] = mu * (s[i] / a[2] /a[2]) ;
 dxda[3] = - mu / (a[3] + s[i]) ;
 for(j=1; j<=na; j++) dxda[j] *= factor ;
}
```

APPENDIX G. COMPUTER PROGRAMS USED IN 3-COMPARTMENT STRUCTURED MODELING

In 3-compartment structured modeling, two computer program were used in order. In programming, subroutines of numerical recipies were used as source files (Press et al.,1988)

```
/* Structured modeling program */
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <nr.h>
#include <nrutil.h>
void mrqminfx();
void mrqcoffx();
#include <nrutil.c>
#include <gaussj.c>
#include <covsrt.c>
#include <mrqminfx.c>
#include <mrqcoffx.c>
#include <ttest.c>
#include <avevar.c>
#include <ftest.c>
#include <betai.c>
#include <betacf.c>
#include <gammln.c>
void eta_of_xg();
float stddev();
float *s, *S, *Xp, *eta, *t, *Xa, *Xt, *Xg;
```

```
main(argc,argv)
int argc;
char *argv[];
{
 int i,j,isum,nsteps,*count,binary,i_fit_func,order,iswap=0,i_imp,i_prec;
 int ndata,ma,mfit,*lista,Sm,index,i_initdata;
 float
*etanew,*sig,*a,*p,**covar,**alpha,chisq,alamda,yfit,*detada,*a_last,*a_init,etamod;
 float tt,f,tprob,fprob;
 char data_S[80],data_Xt[80],data_Xa[80],data_Xp[80],dataout[100],initdata[20];
 float *ycoord, *xcoord, *ynew, min, max, stepsize,base,dev;
 FILE *in,*out;
 void (*fit_func)();
 fit_func = eta_of_xg ;
 if(argc!=8){
  printf("\nenter:  flnm data_Xt data_Xa data_Xp data_S initdata importance base");
  printf("\ndata_Xt is *** Xt versus t *** file");
  printf("\ndata_Xa is *** Xa versus t *** file");
  printf("\ndata_Xp is *** Xp versus t *** file");
  printf("\ndata_S  is *** S  versus t *** file");
  printf("\n\nenter initdata as [1] for Nielsen");
  printf("\n                  [2] for Kazan");
  printf("\n\nenter importance [1] : for high values");
  printf("\n                 [2] : for difference from center");
  printf("\n                 [3] : for difference from center (end point 0 variance)");
  printf("\n                 [4] : for difference from center (first point 0 variance)");
  printf("\nbase is variance power base");
  printf("\n");
  exit(1);
 }
 i_prec=3;
 ndata=120;
 strcpy(data_Xt,argv[1]);
```

```
  strcpy(data_Xa,argv[2]);
  strcpy(data_Xp,argv[3]);
  strcpy(data_S,argv[4]);
  ma = 5;
  i_initdata = atoi(argv[5]);
  i_imp = atoi(argv[6]);
  base = atof(argv[7]);
  mfit = ma ;
  s = vector(0,ndata+1) ;
  S = vector(0,ndata+1) ;
  Xp = vector(0,ndata+1) ;
  Xt = vector(0,ndata+1) ;
  Xg = vector(0,ndata+1) ;
  eta = vector(0,ndata+1) ;
  t = vector(0,ndata+1) ;
  Xa = vector(0,ndata+1) ;
  etanew = vector(0,ndata+1) ;
  detada = vector(1,ma);
  sig = vector(0,ndata+1);
  a = vector(1,ma);
  a_last = vector(1,ma);
  a_init = vector(1,ma);
  p = vector(1,ma);
  lista = ivector(1,ma);
  covar = matrix(1,ma,1,ma);
  alpha = matrix(1,ma,1,ma);
  if(i_initdata==1) sprintf(initdata,"init_Nielsen");
  else if(i_initdata==2) sprintf(initdata,"init_Kazan");
    if((in=fopen(initdata,"rt"))==NULL)exit(1);
    for(i=1; i<=ma; i++) fscanf(in,"%g",&a[i]);
    fclose(in);
 i=0;
  if((in=fopen(data_Xt,"rt"))==NULL)exit(1);
```

```
 while(!feof(in)){
  i++;
  fscanf(in,"%g %g",&t[i],&Xt[i]);
}
 fclose(in);
 if((in=fopen(data_Xp,"rt"))==NULL)exit(1);
 for(i=1; i<=ndata; i++){
  fscanf(in,"%g %g",&t[i],&Xp[i]);
if(index==0){
  for(i=1; i<=ndata; i++){
    eta_of_xg(i,a,&etamod,detada,ma);
    etanew[i] = etamod ;
   }
   sprintf(dataout,"eta_fitinit");
   if((out=fopen(dataout,"wt"))==NULL)exit(1);
   for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",s[i],etanew[i]);
   fclose(out);
   exit(1);
 }
```

/*Second program to calculate X_G and X_A*/

```
#include <math.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <nr.h>
#include <nrutil.h>
void mrqminfx();
void mrqcoffx();
#include <nrutil.c>
#include <gaussj.c>
#include <covsrt.c>
#include <mrqminfx.c>
```

```
#include <mrqcoffx.c>
#include <ttest.c>
#include <avevar.c>
#include <ftest.c>
#include <betai.c>
#include <betacf.c>
#include <gammln.c>
void xg();
void xa();
float stddev();
float *s, *S, *Xp, *eta, *t, *Xa, *Xt, *Xg;
main(argc,argv)
int argc;
char *argv[];
{
 int i,j,isum,nsteps,*count,binary,i_fit_func,order,iswap=0,i_imp,i_prec,idir;
 int ndata,ma,mfit,*lista,Sm,index,i_initdata;
 float
*x_new,*sig,*a,*p,**covar,**alpha,chisq,alamda,yfit,*detada,*a_last,*a_init,etamod;
 float tt,f,tprob,fprob;
 char data_S[80],data_Xt[80],data_Xa[80],data_Xp[80],dataout[100],initdata[20];
 float *ycoord, *xcoord, *ynew, min, max, stepsize,base,dev;
 FILE *in,*out;
 void (*fit_func)();
 fit_func = xg ;
 if(argc!=6){
  printf("\nenter:  flnm data_Xt data_Xa data_Xp data_S initdata");
  printf("\ndata_Xt is *** Xt versus t *** file");
  printf("\ndata_Xa is *** Xa versus t *** file");
  printf("\ndata_Xp is *** Xp versus t *** file");
  printf("\ndata_S  is *** S  versus t *** file");
  printf("\n\n initdata is the parameter file");
  printf("\n");
```

```
  exit(1);
 }
 i_prec=3;
 ndata=120;
 strcpy(data_Xt,argv[1]);
 strcpy(data_Xa,argv[2]);
 strcpy(data_Xp,argv[3]);
 strcpy(data_S,argv[4]);
 strcpy(initdata,argv[5]);
 ma = 5;
 s = vector(0,ndata+1) ;
 S = vector(0,ndata+1) ;
 Xp = vector(0,ndata+1) ;
 Xt = vector(0,ndata+1) ;
 Xg = vector(0,ndata+1) ;
 eta = vector(0,ndata+1) ;
 t = vector(0,ndata+1) ;
 Xa = vector(0,ndata+1) ;
 x_new = vector(0,ndata+1) ;
 detada = vector(1,ma);
 sig = vector(0,ndata+1);
 a = vector(1,ma);
 a_last = vector(1,ma);
 a_init = vector(1,ma);
 p = vector(1,ma);
 lista = ivector(1,ma);
 covar = matrix(1,ma,1,ma);
 alpha = matrix(1,ma,1,ma);
strcpy(initdata,argv[5]);
  if((in=fopen(initdata,"rt"))==NULL)exit(1);
  for(i=1; i<=ma; i++){
    fscanf(in,"\ta[%d]=%g",&i,&a_last[i]);
    printf("\ta[%d]=%g",i,a_last[i]);
```

```
    }
    fclose(in);
i=0;
  if((in=fopen(data_Xt,"rt"))==NULL)exit(1);
  while(!feof(in)){
    i++;
    fscanf(in,"%g %g",&t[i],&Xt[i]);
}
  fclose(in);
  ndata=i-1;
  printf("\ni=%d\tx=%lg",ndata,Xt[ndata]);
  if((in=fopen(data_Xa,"rt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++){
    fscanf(in,"%g %g",&t[i],&Xa[i]);
}
  fclose(in);
  if((in=fopen(data_Xp,"rt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++){
    fscanf(in,"%g %g",&t[i],&Xp[i]);
}
  fclose(in);
  if((in=fopen(data_S,"rt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++){
    fscanf(in,"%g %g",&t[i],&s[i]);
  }
  fclose(in);
  for(i=1; i<=ndata; i++) Xg[i] = Xt[i] - (Xa[i] + Xp[i]) ;
  if((out=fopen("Xg_long","wt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++){
    fprintf(out,"%g %g\n",t[i],Xg[i]);
  }
  fclose(out);
  t[0] = (2*t[1]) - t[2] ; t[ndata+1] = (2*t[ndata]) - t[ndata-1];
```

```
  Xp[0] = (2*Xp[1]) - Xp[2] ; Xp[ndata+1] = (2*Xp[ndata]) - Xp[ndata-1];
  Xa[0] = (2*Xa[1]) - Xa[2] ; Xa[ndata+1] = (2*Xa[ndata]) - Xa[ndata-1];
  Xg[0] = (2*Xg[1]) - Xg[2] ; Xg[ndata+1] = (2*Xg[ndata]) - Xg[ndata-1];
for(i=1; i<=ndata; i++){
    eta[i] = (Xg[i+1] - Xg[i-1]) / (t[i+1] - t[i-1]);
  }
/***************** forward quadrature for Xg ****************************/
  x_new[0] = Xg[0]; x_new[1] = Xg[1] ;
  idir = 1;
  for(i=1; i<=ndata; i++) xg(i,a_last,idir,x_new) ;
dev = stddev(Xg,x_new,ndata);
  ttest(Xg,ndata,x_new,ndata,&tt,&tprob);
  ftest(Xg,ndata,x_new,ndata,&f,&fprob);
  sprintf(dataout,"Xg_%s_idir1",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
  fclose(out);
  sprintf(dataout,"Xg_%s_idir1_stat",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ma; i++)fprintf(out,"\ta[%d]=%g",i,a_last[i]);
  fprintf(out,"\nstddev=%g",dev);
  fprintf(out,"\nt=%g\ttprob=%g",tt,tprob);
  fprintf(out,"\nf=%g\tfprob=%g",f,fprob);
  fclose(out);
/***************** backward quadrature for Xg ****************************/
  x_new[ndata] = Xg[ndata]; x_new[ndata+1] = Xg[ndata+1] ;
  idir = 0;
  for(i=ndata; i>=1; i--) xg(i,a_last,idir,x_new) ;
  dev = stddev(Xg,x_new,ndata);
  ttest(Xg,ndata,x_new,ndata,&tt,&tprob);
  ftest(Xg,ndata,x_new,ndata,&f,&fprob);
  sprintf(dataout,"Xg_%s_idir0",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
```

```
  for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
  fclose(out);
  sprintf(dataout,"Xg_%s_idir0_stat",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ma; i++)fprintf(out,"\ta[%d]=%g",i,a_last[i]);
  fprintf(out,"\nstddev=%g",dev);
  fprintf(out,"\nt=%g\ttprob=%g",tt,tprob);
  fprintf(out,"\nf=%g\tfprob=%g",f,fprob);
  fclose(out);
/**************** forward quadrature for Xa ***************************/
  x_new[0] = Xa[0]; x_new[1] = Xa[1] ;
  idir = 1;
  for(i=1; i<=ndata; i++) xa(i,a_last,idir,x_new) ;
  dev = stddev(Xa,x_new,ndata);
  ttest(Xa,ndata,x_new,ndata,&tt,&tprob);
  ftest(Xa,ndata,x_new,ndata,&f,&fprob);
  sprintf(dataout,"Xa_%s_idir1",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
  fclose(out);
  sprintf(dataout,"Xa_%s_idir1_stat",argv[5]);
  if((out=fopen(dataout,"wt"))==NULL)exit(1);
  for(i=1; i<=ma; i++)fprintf(out,"\ta[%d]=%g",i,a_last[i]);
  fprintf(out,"\nstddev=%g",dev);
  fprintf(out,"\nt=%g\ttprob=%g",tt,tprob);
  fprintf(out,"\nf=%g\tfprob=%g",f,fprob);
  fclose(out);
/**************** backward quadrature for Xa ***************************/
  x_new[ndata] = Xa[ndata]; x_new[ndata+1] = Xa[ndata+1] ;
  idir = 0;
  for(i=ndata; i>=1; i--) xa(i,a_last,idir,x_new) ;
  dev = stddev(Xa,x_new,ndata);
  ttest(Xa,ndata,x_new,ndata,&tt,&tprob);
```

```
 ftest(Xa,ndata,x_new,ndata,&f,&fprob);
 sprintf(dataout,"Xa_%s_idir0",argv[5]);
 if((out=fopen(dataout,"wt"))==NULL)exit(1);
 for(i=1; i<=ndata; i++) fprintf(out,"%g\t%g\n",t[i],x_new[i]);
 fclose(out);
 sprintf(dataout,"Xa_%s_idir0_stat",argv[5]);
 if((out=fopen(dataout,"wt"))==NULL)exit(1);
 for(i=1; i<=ma; i++)fprintf(out,"\ta[%d]=%g",i,a_last[i]);
 fprintf(out,"\nstddev=%g",dev);
 fprintf(out,"\nt=%g\ttprob=%g",tt,tprob);
 fprintf(out,"\nf=%g\tfprob=%g",f,fprob);
 fclose(out);
}
float stddev(x,y,n)
float x[],y[];
int n;
{
 float var=0.; int i;
 for(i=1; i<=n; i++) var += (y[i] - x[i]) * (y[i] - x[i]) ;
 return (sqrt(var));
}
void xg(i,a,idir,x)
float a[],*x;
int i,idir;
{
 float eta,frac1, frac2, sum1, sum2;
 frac1 = (0.9 * s[i]) / ( (0.9 * s[i]) + a[2] ) ;
 frac2 = (0.9 * s[i]) / ( (0.9 * s[i]) + a[1] ) ;
 sum1 = frac1 * Xa[i] * ( (0.9 * a[4]) + (0.1 * x[i] * (a[4] + a[5])) ) ;
 sum2 = frac2 * Xa[i] * x[i] * a[3] ;
 eta = sum1 - sum2 ;
 if(idir==1) x[i+1] =  (eta * (t[i+1]-t[i-1])) + x[i-1];
 else if(idir==0) x[i-1] = - ( eta * (t[i+1]-t[i-1])) + x[i+1];
```

```
}
void xa(i,a,idir,x)
float a[],*x;
int i,idir;
{
  float eta,frac1, frac2, sum1, sum2;
   frac2 = (0.9 * s[i]) / ( (0.9 * s[i]) + a[2] ) ;
  frac1 = (0.9 * s[i]) / ( (0.9 * s[i]) + a[1] ) ;
  sum1 = frac1 * Xa[i] * (1. - Xa[i]) * a[3] ;
  sum2 = frac2 * Xa[i] * (a[4] + a[5]) * ((1.9 * Xa[i]) - 1.) ;
  eta = sum1 - sum2 ;
  if(idir==1) x[i+1] =  (eta * (t[i+1]-t[i-1])) + x[i-1];
  else if(idir==0) x[i-1] = - ( eta * (t[i+1]-t[i-1])) + x[i+1];
}
```

REFERENCES

Aiken, C. and R. I. Gumport, " Restriction Endonuclease *Rsr*I from Rhodobacter sphaeroides, an Isochizomer of *Eco*RI: Purification and Properties", Nucleic Acids Research, Vol. 16, No. 16, pp. 7901-7916, 1988.

Ataai, M. M. and M. L. Shuler, " A Mathematical Model for Prediction of Plasmid Copy Number and Genetic Stability in *Escherichia Coli*", Biotechnology and Bioengineering, Vol. 30, pp. 389-397, 1987.

Atlas, R. M., *Microbiology Fundemantals and Applications*, Second Edition, Edward Arnold Publishers, London, 1991.

Backman, K., M. Ptashne, "Maximizing gene expression on a plasmid using recombination in vitro", Cell, Vol. 13, pp65-71, 1978.

Backman, K., M. Ptashne and W. Gilbert, " Construction of Plasmids Carrying the cI Gene of Bacteriophage λ", Proc. Natl. Acad., Vol.73, No. 11, pp 4174-4178, 1976.

Bailey, J. E., D. F. Ollis, *Biochemical Engineering Fundemantals*, Second Edition, McGraw-Hill, Singapore, 1986.

Baltes M., R. Schneider, C. Sturm and M. Reuss, " Optimal Experimental Design for Parameter Estimation in Unstructured Growth Models", Biotechnol. Prog., Vol. 10, pp 480-488, 1994.

Bentley, W. E. and O.E. Quiroga, " Investigation of Subpopulation Heterogeneity and Plasmid Stability in Recombinant *Escherichia coli* via a Simple Segregated Model", Biotechnology and Bioengineering, Vol. 42, pp. 222-234, 1993.

Betenbaugh, M. J., C. Beaty and P. Dhurjati, " Effects of Plasmid Amplification and Recombinant Gene Expression on the Growth Kinetics of Recombinant *E.coli*", Biotechnology and Bioengineering, Vol. 33, pp 1425-1436, 1989.

Bingham, A. H. A., A. F. Sharman and T. Atkinson, " The Purification of Restriction endonuclease *Eco*RI by Precipitation Involving Polyethyleneimin", FEBS Letters, Vol. 76, No. 2, pp 250-256, April 1977.

Botterman, J. and M. Zabeau, " High Level Production of the *Eco*RI Endonuclease Under the Control of the p_L Promoter of the Bacteriophage Lambda", Gene, Vol. 37, pp 229-239, 1985.

Botterman, J. H., D.R. De Buyser, J. A. Spriet, M. Zabeau and G. C. Vansteenkiste, "Fermentation and Recovery of the *Eco*RI Restriction Enzyme with a Genetically Modified *Escherichis coli* Strain", Biotechnology and Bioengineering, Vol. 27, pp. 1320-1327, 1985.

Bowerman, O., *Linear Statistical Models: An Applied Approach*, Second Edition, PWS-Kent Publishing Company, Boston, 1990.

Bradford, M. M., " A Rapid and Sensitive Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-dye Binding", Analytical Biochemistry, Vol. 72, pp 284-254, 1976.

Brantl, S. and D. Behnke, " The Amount of RepR Protein Determines the Copy Number of Plasmid pIP501 in *Bacillus subtilis*", Journal of Bacteriology, pp 5475-5478, Aug. 1992.

Broda, P., *Plasmids*, Western Printing Services, Bristol, 1979.

Cazzador, L., " Comments on a Segregated Model of Recombinant Cultures", Biotechnology and Bioengineering, Vol. 37, pp 981-984, 1991.

Cheng, S, R. Kim, K. King, S. Kim and P. Modrich, " Isolation of Gram Quantities of *Eco*RI Restriction and Modification Enzymes from an Overproducing Strain", The Journal of Biological Chemistry, Vol.259, No. 18, pp. 11571-11575, Sept. 1984.

Cooper, N. S., M. Brown and C. A. Caulcott, " A Mathematical Method for Analysising Plasmid Stability in Micro-organisms", Journal of General Microbiology, Vol. 133, pp 1871-1880, 1987.

Coppela, S. J., C. M. Acheson and P Dhurjati, "Measurment of Copy Number Using HPLC", Biochemical and Bioengineering, Vol. 24, pp 646-647, 1987.

Davidson, A. M., A. Dunn, M. J. Day and P. F. Randerson, " A nonlinear Technique fir the Analysis of the Plasmid Instability in Micro-organisms", Journal of General Microbiology, Vol. 136, pp 59-64, 1990.

Dechow, F. J., *Separation and Purification Techniques in Biotechnology*, Noyes Publication, New Jersey, 1989.

diPasquantonio, V. M., M. Betenbaugh and P. Dhurjati, " Improvement of Product Yıelds by Temperature-Shifting of Escherichia coli Cultures Containing Plasmid pOU140", Biotechnology and Bioengineering, Vol. 29, pp 513- 519, 1987.

Dubey, A. K., S. N. Mukhopadhyay, V. S. Bisaria and T. K. Ghose, " Sources, Production and Purification of Restricition Enzymes", Process Biochemistry, pp 25-34, 1987.

Du Poet, P. T., Y. Arcand, R. Bernier, J. N. Barbotin, and D. Thomas, " Plasmid Stability in Immobilized and Free Recombinant *Escherichia coli* JM105 (pKK223-200): Importance of Oxygen Diffusion , Growth Rate and Plasmid Copy Number", Applied and Environmental Microbiology, pp 1548-1555, July 1987.

Emerick, A.W., B.L. Bertalini, A. Ben-Bassal, J.J White, M.W. Konrad, " Expression of a β-prepoinsulin fusion protein in E.coli", Bio. Tech., Vol.2, pp165-168, 1984.

Flaschel, E. and K. Friehs, " Improvement of Downstream Processing of Recombinant Proteins by Means of Genetic Engineering Methods", Biotechnology Advances, Vol. 11, pp 31-78, 1993.

Greene P. J., M. C. Betlach and H. W. Boyer, " The EcoRI Restriction Endonuclease", Methods Mol. Biol., Vol. 7, pp 87-111, 1974.

Greene, P. J., H. L. Heyneker,., F. Bolivar, R. L. Rodriguez, M. C. Betlach, A. A. Covarrubias, K. Backman, D. J. Russel, R. Tait and H. W. Boyer, " A general Method for the Purification of Restriction Enzymes", Nucleic Acids Research, Vol. 5, No. 7, pp.2373-2380, 1978.

Greene, P., M. Gupta, H. W. Boyer, W. E. Brown and J. M. Rosenberg, " Sequence Analysis of the DNA Encoding the *Eco*RI Endonuclease and Methylase", The Journal of Biological Chemistry, Vol. 256, No.5, pp 2143-2153, March 1981.

Hale, W. G. and Margham, J. P, *Biology*, Harper Collins Publishers, New York, 1991.

Halford, S. E. and N. P. Johnson, " The *Eco*RI Endonuclease with Bacteriophage λ DNA", Biochem. J., Vol.191, pp 593-604, 1980.

Han, K. and O. Levenspiel, " Extended Monod Kinetics for Substrate, Product and Cell Inhibition", Biotechnology and Bioengineering, Vol. 32, pp 430-437, 1988.

Harris, E. L. V., S. Angal, *Protein Purification Methods*: A Practical Approach, IRL Press, 1990.

Heitman, J., " How *Eco*RI Endonuclease Recognizes and Cleaves DNA", Bioessays, Vol. 14, No. 7, pp 445-454, July, 1992.

Heitman, J., " On the Origins, Structures and Functions of Restriction -Modification enzymes", Genetic Engineering, Vol. 15, pp 57-108, 1993.

Howe, C., *Gene Cloning and Manipulation*, Cambridge University Press, New York, 1995.

Huang, J., P. Dhulster, D. Thomas and J. N. Barbotin, " Agitation Rate Effects on Plasmid Stability in Immobilized and Free-Cell Continous Cultures of Recombinant *E.coli*", Enzyme Microb. Technol., Vol. 12, pp 933-939, 1990.

Huang, C.T., S. W. Peretti and J. D. Bryers, " Plasmid Retention and Gene Expression in Suspended and Biofilm Cultures of Recombinant *Escherichia coli* DH5α (pMJR1750)", Biotechnology and Bioengineering, Vol. 41, pp 211-219, 1993.

Imanaka, T. and Aiba, S., " A Perspective on the Application of Genetic engineering: Stability of Recombinant Plasmid", Annals New York Academy of Sciences, pp 1-14, 1981.

Iniesta, M., " Kinetic Study of pTG201 Plasmid Stability in *Escherichia coli*", Acta Biotechnol., Vol. 8, pp 311-318, 1988.

Ivanov, I. G., and Bachvarov, D. R. " Determination of Plasmid Copy Number by the Boiling Method", Analytical Biochemistry, Vol. 165, pp 137-145, 1987.

Jack, E.W., L.Greenough, L.F.Domer, Shuang-yong Xu, T. Strzelecka, A. Aggarwal and I. Schildkraut, "Overexpression, purification and crystallization of *Bam*HI endonuclease", Nucleic Acid Research, Vol.19, No.8, pp.1825-1829,1991.

Jeltsch, A., J. Alves, G. Maass and A. Pingoud, " On the Catalytic Mechanism of *Eco*RI and *Eco*RV", FEBS Letters, Vol. 304, No. 1, pp 4-8, 1992.

Kaszubska, W., C. Aiken, C. D. O'Connor and R. I. Gumport, " Purification, Cloning and Sequence Analysis of *Rsr*I DNA Methyltransferase: Lack of Homology Between Two Enzymes, *Rsr*I and *Eco*RI, that Methylate the Same Nucleotide in Identical Recognition Sequence", Nucleic Acids Research, Vol. 17, No. 24,pp.10403-10425, 1989.

Koizumi, J., Y. Monden and S. Aiba, " Effects of Temperature and Dilution Rate on the Copy Number of Recombinant Plasmid in Continuous Culture of Bacillus stearothermophilus (pLP11)", Biotechnology and Bioengineering, Vol. 27, pp. 721-728, 1985.

Kumar, P. K. R., H. E. Maschke, K. Friehs and K. Schügerl, " Strategies for Improving Plasmid Stability in Genetically Modified Bacteria in Bioreactors", TIBTECH, Vol. 9, pp 279-284, 1991.

Laemnli, U. K:, " Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4", Nature, Vol. 227, pp 680-685, 1970.

Laffend, L. and M. L. Shuler, " Structured Model of Genetic Control via the lac Promoter in *Escherichia coli*", Biotechnology and Bioengineering, Vol.43, pp 399-410, 1994.

Lamotte, D., A. Gschaedler and J. Boudrant, " Transitory Plasmid Instability During the Exponential Phase of Growth of a Recombinant *Escherichia coli*", Biotechnology Letters, Vol. 16, No. 11, pp 1113- 1118, 1994.

Lavastida, J., L. Colon, M. Ward and B. C. Stark, " Improvement in Segregational Stability of Recombinant Plasmids by Retransformation of *E.coli* Host Cells", Biotechnology Letters, Vol. 15, No. 8, pp 779-784, 1993.

Lee, J. and W. F. Ramirez, " Mathematical Modeling of Induced Foreign Protein Production by Recombinant Bacteria", Biotechnology and Bioengineering, Vol. 39, pp 635-646, 1992.

Lefebvre, J., G. A. Junter and J. C. Vincent, " Graphics- Associated Modelling of Batch Cultures of *Escherichia coli* Fermenting Glucose", Enzyme Microb. Technol., Vol. 16, pp 163-169, 1994.

Lewington, J. and M. J. Day, "A Rapid Electrophoretic Method for the Measurement of Plasmid Copy Number", Letters in Applied Microbiology, Vol. 3, No.6, pp 109-112, 1986.

Luke P. A. and S. E. Halford, " Solubility of the *Eco*RI Restriction Endonuclease and its Purification from an Overproducing Strain", Gene, Vol. 37, pp. 241-246, 1985.

Luong, J. H. T., " Generalization of Monod Kinetics for Analysis of Growth Data with Substrate Inhibition", Biotechnology and Bioengineering, Vol. 29, pp 242-248, 1987.

Mayer, H., " Optimization of the *Eco*RI* Activity of *Eco*RI Endonuclease", FEBS Letters, Vol. 90, No.2, pp 341-344, June 1978.

Mehra, R. S., V. P. Malhotra, G. W. Rembhotkar, "Rapid Purification of a Restriction Endonuclease From *Escherichia coli* RY13", Biotechnology Techniques, Vol. 7, No. 6, pp 411-414, 1993.

Meyer, H.P., H.J. Kuhn, S.W. Brown, A. Fiecher, "Production of humanleucocylce interferon by *E.coli*", Proc.3rd Eur. Comg. Biotechnol., Vol.1, pp499-505, 1984

Modrich P. and D. Zabel, " *Eco*RI Endonuclease: Physical and Catalytic Properties of the Homogeneous Enzyme", The Journal of Biological Chemistry, Vol. 251, No. 19, pp 5866-5874, 1976.

Naveh, D., " Scale-Up of Fermentation for Recombinant DNA Products", Food Technology, pp 102-109, 1985.

Needles M.C.,S.R. Fried, R. Love, J.M. Rosenberg, H.W. Boyer and P.J. Greene, "Determinants of *Eco*RI endonuclease sequence discrimination", Proc. Natl. Acad. Sci., Vol. 86, pp 3579-3583, 1989.

Neubauer, P., K.Hofmann, O. Holst, B. Mattiasson and P. Krusche, " Maximizing the Expression of a Recombinant Gene in *Escherichia coli* by Manipulation of Induction Time Using Lactose as Inducer", Applied Microbiology Biotechnology, Vol. 36, pp. 739-744, 1992.

Nielsen, J., A. G. Pedersen, K. Strudsholm and J. Villadsen, " Modeling Fermentations with Recombinant Microorganisms: Formulation of a Structured Model ", Biotechnology and Bioengineering, Vol. 37, pp 802-808, 1991a.

Nielsen, J., C. Emborg, K. Halberg and J. Villadsen, " Compartment Model Concept Used in the Design of Fermentation with Recombinant Micro-organisms", Biotechnology and Bioengineering, Vol. 34, pp 478-486, 1989.

Nielsen, J., K. Nikolajsen and J. Villadsen, "Structured Modeling of a Microbial Systems: A Theoretical study of Lactic Acid Fermentation", Biotechnology and Bioengineering, Vol. 38, pp 1-10, 1991b.

Nielsen, J., K. Nikolajsen and J. Villadsen, "Structured Modeling of a Microbial Systems Experimental Verification of a Structured Lactic Acid Fermentation Model", Biotechnology and Bioengineering, Vol. 38, pp 11-23, 1991c.

Nikolajsen, K., J. Nielsen, and J. Villadsen, "Structured Modeling of a Microbial Systems: III. Growth on Mixed Substrates", Biotechnology and Bioengineering, Vol. 38, pp 24-29, 1991.

Old, R. W., S. B. Primrose, *Principle of Gene Manipulation*, Fourth Edition, Blackwell Scientific Publications, Bilbao, 1989.

Perbal, B. , *A Practcal Guide to Molecular Cloning*, Second Edition, John Wiley & Sons, New York, 1988.

Peretti, S. W. and J. E. Bailey, " Mechanistically Detailed Model of Cellular Metabolism for Glucose-Limited Growth of *Escherichia coli* B/r-A", Biotechnology and Bioengineering, Vol. 28, pp 1672-1689, 1986.

Pirrotta,V. and T. A. Bickle, " General Purification Schemes for Restriction Endonucleases", Methods in Enzymology, Vol. 65, pp 89-95, 1980.

Pirt, S. J., *Principles of Microbe and Cell Cultivation*, Backwell Scientific Publication, London, 1973.

Press, W. H., B.P. Flannery, S.A. Teukolsky, W. T. Vetterling, *Numerical Recipies in C*, Cambridge University Press, New York, 1988.

Primrose, S. B., *Molecular Biotechnology*, Blackwell Scientific Publications, London, 1991.

Primrose S.B, S.D. Ehrlich, " A perspective on the application of genetic engineering: Stability of recombinant plasmid", Vol.16,pp 193-201, 1981.

Rehm, H. J., G. Reed, A. Pühler, P. Stader, *Biotechnology-Fundamentals of Biochemical Engineering*, Vol. 1,Second Edition, Berlin, 1991.

Rehm, H. J., G. Reed, H. Brauer (volume editor), *Biotechnology-Fundamentals of Biochemical Engineering*, Vol. 2, Berlin, 1985.

Roberts, R. J. and D. Macelis, " Restriction Enzymes and Their Isochizomers", Nucleic Acids Research, Vol. 19, Supplement, pp 2077-2109, 1991.

Roels J. A., *Energetics and Kinetics in Biotechnology*, Elsevier Biomedical Press, Amsterdam, 1983.

Rosenberg, J. M., " Structure and Function of Restriction Endonucleases", Current Opinion in Structural Biology, Vol. 1, pp 104-113, 1991.

Rubin, R.A. and P. Modrich, " *Eco*RI Methylase: Physical and Catalytic Properties of the Homogeneous Enzyme", The Journal of Biological Chemistry, Vol. 252, No. 20, pp 7265-7272, 1977.

Rubin, R.A. and P. Modrich, " Purification and Properties of *Eco*RI Endonuclease", Methods in Enzymology, Vol. 65, pp 96-104, 1980.

Ryan, W. and S. J. Parulekar, " Recombinant Protein Synthesis and Plasmid Instabilty in Continous Cultures of Escherichia coli JM103 Harboring a High Copy Number Plasmid", Biotechnology and Bioengineering, Vol. 37, pp 415-429, 1991.

Ryu, D.D. and S. B. Lee, " Development and Optimization of Recombinant Fermentation Processes", Horizon of Biochemical Engineering, pp 97-124, 1987.

Sambrook, J., E. F. Fritsch and T. Maniatis, *Molecular Cloning-A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, New York, 1989.

Schendel, F. J., E. J. Baude and M. C. Flickenger, " Determination of Protein Expression and Plasmid Copy Number from Cloned Genes in Escherichia coli by Flow Injection Analysis Using an Enzyme Indicator Vector", Biotechnolgy and Bioengineering, Vol. 34, pp 1023-1036, 1989.

Seo J. and J. E. Bailey, " A Segregated Model for Plasmid Content and Product Synthesis in Unstable Binary Fission Recombinant Organisms", Biotechnology and Bioengineering, Vol. 27, pp 156- 165, 1985.

Schildkraut, I. , " Screening for and Characterizing Restriction Endonucleases", Genetic Engineering, Vol. 6, pp 117-140, 1993.

Shimatake, H. and M. Rosenberg, " Purified l Regulatory Protein cII Positively Activates Promoters for Lysogenic Development", Nature, Vol. 292, No.9, July, 1981.

Shu, J. and M. L. Shuler, " Amino Acid Supplementation Decreases Plasmid Retention in *Escherichia coli*", Biotechnology and Bioengineering, Vol. 40, pp. 1197-1202, 1992.

Smith, C. A., J. Wood, *Molecular Biology and Biotechnology*, Chapman and Hall, Hong Kong, 1991.

Sümegi, J., D. Breedveld, P. Hossenloop and P. Chambon, " A Rapid Procedure for Purification of *Eco*RI Endonuclease", Biochemical and Biophysical Research Communications, Vol. 76, No.1, pp 78-85, 1977.

Terry, B. J., Jack, W. E., Rubin, R. A. and Modrich P., " Thermodynamic Parameters Governing Interaction of *Eco*RI Endonuclease with Specific and Nonspecific DNA Sequences", The Journal of Biological Chemistry, Vol. 258, No 16, pp 9820-9825, August 1983.

Vila, P., J. L. Corchero, A. Benito and A. Villaverde, " Ammonium- Mediated Reduction of Plasmid Copy Number and Recombinant Gene Expression in *Escherichia coli*", Biotechnol. Prog., Vol. 10, pp 648-651, 1994.

Vlatakis, G. and V. Bouriotis, " Sequence-Specific DNA Affinity Chromatography: Application to the Purification of *Eco*RI and *Sph*I", Methods in Enzymology, pp 352-358, 1991.

Weathers, P., " Biochemistry in an Industrial Context: Methods of Protein Purification and Downstream Processing", Journal Chemical Education, Vol. 65, No. 10, pp 854-856, 1988.

Wilson, K. K., K. H. Goulding, *A Biologist's Guide to principles and Techniques of Practical Biochemistry*, ELBS, London, 1991.

Winkler, F. K., " Structure and Function of Restriction Endonucleases", Current Opinion in Structural Biology, Vol. 2, pp 93-99, 1992.

Wu, X., L. Brandes, H. E. Maschke and K. Schügerl, " Overproduction of *Eco*RI and *Sp*A: *Eco*RI Fusion Protein by *Escherichia coli* Employing the Combinations of Three Different Types of Multicopy Plasmids: Comparison of Plasmid Stabilities and Product Concentrations", J. Chem. Tech. Biotechnol., Vol 56, pp 383-387, 1993.