BIOPROCESS DEVELOPMENT FOR THERAPEUTICAL PROTEIN PRODUCTION

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EDA ÇELİK AKDUR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
CHEMICAL ENGINEERING

DECEMBER 2008

Approval of the thesis:

BIOPROCESS DEVELOPMENT FOR THERAPEUTICAL PROTEIN PRODUCTION

submitted by **EDA ÇELİK AKDUR** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Chemical Engineering Department, Middle East Technical University** by,

Prof.Dr. Canan Özgen _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof.Dr. Gürkan Karakaş _____
Head of Department, **Chemical Engineering**

Prof.Dr. Pınar Çalık _____
Supervisor, **Chemical Engineering Dept., METU**

Prof.Dr. Stephen G. Oliver _____
Co-Supervisor, **Biochem. Dept., University of Cambridge, UK**

Examining Committee Members:

Prof.Dr. Tunçer H. Özdamar _____
Chemical Engineering Dept., Ankara University

Prof.Dr. Pınar Çalık _____
Chemical Engineering Dept., METU

Prof.Dr. İsmail Tosun _____
Chemical Engineering Dept., METU

Prof.Dr. Tülin Kutsal _____
Chemical Engineering Dept., Hacettepe University

Prof.Dr. Semra Kocabıyık _____
Biology Dept., METU

Date: 19. 12. 2008

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Eda Çelik Akdur

Signature :

ABSTRACT

BIOPROCESS DEVELOPMENT FOR THERAPEUTICAL PROTEIN PRODUCTION

Çelik Akdur, Eda
Ph.D., Department of Chemical Engineering
Supervisor: Prof.Dr. Pınar Çalık
Co-Supervisor: Prof.Dr. Stephen G. Oliver

December 2008, 270 pages

In this study, it was aimed to develop a bioprocess using the *Pichia pastoris* expression system as an alternative to the mammalian system used in industry, for production of the therapeutically important glycoprotein, erythropoietin, and to form stoichiometric and kinetic models.

Firstly, the human EPO gene, fused with a polyhistidine-tag and factor-Xa protease target site, in which cleavage produces the native termini of EPO, was integrated to *AOX1* locus of *P. pastoris*. The Mut^+ strain having the highest rHuEPO production capacity was selected. The glycosylation profile of rHuEPO was characterized by MALDI-ToF MS and Western blotting. The native polypeptide form of human EPO was obtained for the first time in *P. pastoris* expression system, after affinity-purification, deglycosylation and factor-Xa protease digestion.

Thereafter, effects of medium components and pH on rHuEPO production and cell growth were investigated in laboratory-scale bioreactors. Sorbitol was shown to increase production efficiency when added as a co-substrate. Moreover, a cheap alternative nutrient, the byproduct of biodiesel industry, crude-glycerol, was suggested for the first time for *P. pastoris* fermentations. Furthermore, methanol feeding strategy was investigated in fed-batch pilot-scale bioreactors, producing 70 g L^{-1} biomass and 130 mg L^{-1} rHuEPO at t=24h.

Moreover, metabolic flux analysis by using the stoichiometric model formed, which consisted of m=102 metabolites and n=141 reactions, proved useful in further understanding the *P. pastoris* metabolism. Finally, the first structured kinetic model formed for r-protein production with *P. pastoris* successfully predicted cell growth, substrate consumption and r-product production rates, where rHuEPO production kinetics was associated with AOX production and proteolytic degradation.

Keywords: Recombinant human erythropoietin, *Pichia pastoris*, fermentation optimization, metabolic flux analysis, kinetic modeling

öz

TERAPATİK PROTEİN ÜRETİMİ İÇİN BİYOPROSES GELİŞTİRİLMESİ

Çelik Akdur, Eda
Doktora, Kimya Mühendisliği
Tez Yöneticisi: Prof.Dr. Pınar Çalık
Ortak Tez Yöneticisi: Prof.Dr. Stephen G. Oliver

Aralık 2008, 270 sayfa

Bu çalışmada, terapatik proteinlerden glikoprotein yapısındaki eritropoietin (EPO) hormonu üretimi için, endüstride kullanılan memeli hayvan hücreleri yerine, rekombinant *Pichia pastoris* mayası geliştirilmesi ve geliştirilen mikroorganizma kullanılarak bir biyoproses tasarlanması, ayrıca stokiyometrik ve kinetik modellerinin kurulması amaçlanmıştır.

İnsan EPO hormonun üretim ortamından saflaştırılmasını sağlayan 6xhistidin dizinini ve saflaştırılmış hormonun 6xhistidin dizininden ayrılmasını sağlayan faktör-Xa proteaz enziminin tanıdığı amino asit dizinini kodlayan DNA, EPO genine entegre edilmiş ve oluşturulan *pPICZaA::epo*, *P. pastoris* kromozomunun *AOX1* bölgesine entegre edilmiştir. Doğru klonlar tespit edildikten sonra, üretim kapasitesi en yüksek olan koloni seçilmiştir. Üretilen rHuEPO'nun MALDI-ToF MS ve Western blot teknikleri kullanılarak yapısal analizi gerçekleştirilmiştir. Afinite saflaştırma, deglikozilleme ve faktör-Xa proteaz ile kesme işlemlerinden sonra, 18 kDa büyüklüğündeki EPO öz polipeptidi, *P. pastoris* hücresinden ilk kez elde edilmiştir.

Sonraki aşamalarda, ortam bileşenlerinin ve pH'ın rHuEPO üretimi ve hücre çoğalması üzerine etkileri laboratuvar ölçekli biyoreaktörlerde incelenmiştir. *P. pastoris* ile yarı-kesikli biyoreaktörlerde, ikincil karbon kaynağı olarak sorbitolün ortama eklenmesinin üretimi arttırdığı bulunmuştur. Ayrıca, *P. pastoris* fermentasyon prosesleri için ucuz bir karbon kaynağı olarak, biyodizel endüstrisinin yan ürünü olan ham gliserolün kullanılması, ilk kez bu çalışmada önerilmiştir. Karbon kaynağı besleme stratejisi, yarı kesikli biyoreaktörlerde incelenmiş, t= 24 st'te 70 g L^{-1} kuru hücre ve 130 mg L^{-1} rHuEPO üretilmiştir.

M=102 metabolit ve n=141 reaksiyondan oluşan stokiyometrik model kullanılarak gerçekleştirilen metabolik akı analizi, *P. pastoris*'in metabolizmasında metanol ve metanol+sorbitolün tepkime ağına etkisini aydınlatmıştır. *P. pastoris* ile rekombinant protein üretiminin ilk kez oluşturulan, rHuEPO üretim kinetiğini alkol oksidaz ve proteaz üretimleri ile bağdaştıran yapısal kinetik modeli, hücre çoğalması, substrat tüketimi ve r-protein üretimini doğru olarak saptamıştır.

Anahtar Kelimeler: Rekombinant eritropoietin hormonu, *Pichia pastoris*, fermentasyon optimizasyonu, metabolik akı analizi, kinetik modelleme

To my family

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my supervisor Prof.Dr. Pınar Çalık for her continuous support, guidance and help, in all the possible way, throughout this study. Without her encouragements and practical solutions to problems, I would not have achieved this project with such success.

I would like to specially thank my co-supervior Prof.Dr. Stephen G. Oliver for accepting me in his laboratories in University of Manchester for a year, and sharing his invaluable time and knowledge with me. I have learned a lot in his group.

I am also thankful to Prof.Dr. Tunçer H. Özdamar and Prof.Dr. İsmail Tosun for accepting to be a member of my progress jury and for all their advices and critics, which made this study have more deepness.

I would like to express my thanks to Dr. S. Mitchell Halloran for carrying out MALDI-ToF mass spectrometry analysis and willing to share all his knowledge on MALDI-ToF MS with me.

I would like to thank to all academic, administrative and technical staff of Department of Chemical Engineering, METU, for their help and support throughout my education.

This work was supported by the Middle East Technical University Research Fund Project (BAP -2006-07-02-00-01), Scientific and Technical Research Council of Turkey (TÜBİTAK-MAG 107M258) and the infrastructure used in the experiments by SPO (Turkey) Grants 2001K121030.

My national and international PhD scholarships provided by Scientific and Technical Research Council of Turkey (TUBITAK-BİDEB; 2211 and NATO-A2) are gratefully acknowledged.

Technical services provided by the Biomolecular Analysis Facility and the Sequencing Facility at The University of Manchester and by the Biotechnology Institute at Ankara University, are also acknowledged.

I am grateful to countless number of friends, who were beyond being colleagues. Firstly, to our research group in Industrial Biotechnology and Metabolic Engineering Laboratory, for their support, great friendship, advice and encouragement throughout my studies. As I have been one of the earliest comers to the group, I have met many wonderful people, who are still in our group or continuing their education/career somewhere else. I am also grateful to all the members of other research groups; namely the Biochemical Reaction Engineering Laboratory in Ankara University, the "Yeasties" formerly in University of Manchester (now in Cambridge University) and the research group of Prof. Dr. Ufuk Bakir. I am lucky to have met Dr. Pınar Pir at University of Manchester, who was a post-doc in the "Yeasties" group. Last here, but not the least, many friends I have met in this Department, you made my days more colorful in many aspects. I think they all know who they are.

Above all, I would like to deeply thank to my family and my husband, for loving, supporting and encouraging me all through my life. Thanks for existing at all.

TABLE OF CONTENTS

ABSTRACT iv

ÖZ vi

ACKNOWLEDGEMENTS ix

TABLE OF CONTENTS xi

LIST OF TABLES xvii

LIST OF FIGURES xix

NOMENCLATURE xxiv

CHAPTERS

1. INTRODUCTION 1

2. LITERATURE SURVEY 5

2.1 Bioprocess Development 5

2.1.1 Product: Erythropoietin (EPO) 6

2.1.2 Microorganism Selection 7

2.1.2.1 *Pichia pastoris* 9

2.1.2.1.1 General Characteristics 9

2.1.2.1.2 Glycerol, Methanol and Sorbitol Metabolisms 9

2.1.2.1.3 Expression with Alternative Promoters 13

2.1.2.1.4 Secretion of Proteins and Post-Translational Modifications 14

2.1.2.2 Genetic Engineering of the Microorganism: Techniques and Methodology 17

2.1.3 Medium Design and Bioreactor Operation Parameters 22

2.1.4 Recombinant Protein Purification 30

2.1.5 Structural Analysis of the Recombinant Protein 32

2.1.5.1 SDS-Polyacrylamide Gel Electrophoresis 32

2.1.5.2 Western Blotting 33

2.1.5.3 MALDI-ToF Mass Spectrometry 33

2.2 Computation of Bioprocess Characteristics 35

2.2.1 Yield Coefficients and Specific Rates 35
2.2.2 Oxygen Transfer Characteristics 40
2.3 Mathematical Modelling of Metabolism 44
2.3.1 Stoichiometric Models 44
2.3.1.1 Major Metabolic Pathways of Yeasts and Development of the Intracellular Biochemical Reaction Network 48
2.3.2 Kinetic Models 51
3. MATERIALS AND METHODS 55
3.1 Chemicals 55
3.2 Laboratory Equipment 55
3.3 Buffers and Stock Solutions 55
3.4 Strains, Plasmids and Maintenance of Microorganisms 55
3.5 Growth Media 56
3.6 Genetic Engineering Techniques 57
3.6.1 Plasmid DNA Isolation from *E. coli* 57
3.6.2 Agarose Gel Electrophoresis 57
3.6.3 DNA Extraction from Agarose Gels 57
3.6.4 Primer Design 58
3.6.5 Polymerase Chain Reaction (PCR) 59
3.6.5.1. DNA Purification after PCR 59
3.6.6 Digestion of DNA using Restriction Endonucleases 59
3.6.6.1. DNA Purification after Digestion 60
3.6.7 Ligation 60
3.6.8 Transformation of *E. coli* 61
3.6.9 DNA Sequencing 61
3.6.10 Transfection of *Pichia pastoris* 62
3.6.11 Isolation of Genomic DNA from Yeast 63
3.6.12 Southern Blotting 64
3.7 Recombinant Protein Production and Purification 65
3.7.1 Recombinant Protein Production 65
3.7.1.1 Precultivation 65
3.7.1.2 RHuEPO Production in Laboratory Scale Air Filtered Shake Bioreactors 65
3.7.1.3 RHuEPO Production in the Pilot Scale Bioreactor 65
3.7.1.4 Recombinant Protein Production using Biodiesel Byproduct Crude Glycerol 66

3.7.2 Ultrafiltration....68
3.7.3 Polyhistidine-Tag Affinity Purification of rHuEPO....68
3.7.4 Factor Xa Digestion of rHuEPO....68
3.7.5 Deglycosylation of rHuEPO....69
3.8 Analysis....69
3.8.1 Cell Concentration....69
3.8.2 Total Protein Concentration....69
3.8.3 RHuEPO Concentration....69
3.8.4 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)....70
3.8.4.1 Staining the SDS-PAGE Gels....70
3.8.5 Western Blotting....70
3.8.6 Mass Spectrometry....71
3.8.7 Glycerol, Methanol and Sorbitol Concentrations....72
3.8.8 Amino Acids Concentrations....72
3.8.9 Organic Acids Concentrations....73
3.8.10 Protease Activity Assay....73
3.8.11 AOX Activity Assay....74
3.8.12 Liquid Phase Mass Transfer Coefficient and Oxygen Uptake Rate...75
3.9 Mathematical Modeling....75
3.9.1 The Stoichiometric Model and Mass Flux Balance-Based Analysis....75
3.9.2 The Kinetic Model....75
4. RESULTS AND DISCUSSION....77
4.1 Development of the Recombinant *Pichia pastoris* Producing RHuEPO....77
4.1.1 Propagation and Purification of pPICZαA....79
4.1.2 Amplification of EPO cDNA from *E. coli* IOH44362 by PCR....79
4.1.3 Digestion and Ligation Reactions....80
4.1.4 Transformation of *E. coli* cells with pPICZαA::*epo* and Selection of the True Transformant....82
4.1.5 Transfection of *P. pastoris* cells with pPICZαA::*epo*....84
4.1.6 PCR and Southern Blotting Controls to Select the True Transformants....86
4.2 Expression and Purification of Recombinant Human Erythropoietin in *Pichia pastoris* for Structural Analysis....89
4.2.1 Western Blot Analysis for the Selection of Recombinant *P. pastoris* Strains with the Highest rHuEPO Production Capacity....89

4.2.2 Q-ToF Mass Spectrometry Analysis for Confirmation of Primary Structure of rHuEPO 91
4.2.3 Purification and Obtaining the Native rHuEPO 93
4.2.3.1 Ultrafiltration 93
4.2.3.2 Polyhistidine–tag Affinity Purification 94
4.2.3.3 Deglycosylation and Factor Xa Protease Digestion to Obtain the Native Protein 96
4.2.4 Structural Analysis by MALDI-ToF MS 99
4.2.4.1 Initial Optimization Studies 99
4.2.4.2 The Main Results 101
4.3 Expression of Recombinant Human Erythropoietin in *P. pastoris* for Optimization Studies in Laboratory Scale Air Filtered Shake Bioreactors 106
4.3.1 Complex Medium Design 107
4.3.1.1 Glycerol from Biodiesel via Transesterification of Various Vegetable Oils 107
4.3.1.1.1 Preparation of the Glycerol Phase of the Transesterification Reaction for Fermentation Process 108
4.3.1.2 The Bioprocess by *P. pastoris* for Recombinant Protein Production using Crude Glycerol 109
4.3.1.2.1 Effect of Biodiesel Byproduct Glycerol on Substrate Consumption Profiles 109
4.3.1.2.2 Effect of Biodiesel ByProduct Glycerol on Cell Growth Profiles 110
4.3.1.2.3 Effect of Biodiesel Byproduct Glycerol on Recombinant Protein Production 112
4.3.1.2.4 Yields Coefficients 114
4.3.2 Defined Medium Design 116
4.3.2.1 Effects Macro and Micronutrients on Recombinant *P. pastoris* Growth 116
4.3.2.2 Effect of Initial Sorbitol Concentration on Recombinant *P. pastoris* Growth and RHuEPO Production 118
4.3.2.3 Effect of Methanol Induction Strategy on Recombinant *P. pastoris* Growth 120
4.3.2.4 The non-Inhibitory Sorbitol Concentration 122
4.3.3 Bioreactor Operation Parameters 122
4.3.3.1 Effect of pH on RHuEPO Production and Stability 124

4.4 Expression of Recombinant Human Erythropoietin by *P. pastoris* in Pilot Scale Bioreactors 126
4.4.1 Control of Bioreactor Operation Parameters in Pilot-Scale Bioreactor 126
4.4.2 Effect of Sorbitol and Methanol Feeding Rate in Fed-Batch Pilot Scale Bioreactor Operations 128
4.4.2.1 Effect of Sorbitol and Feeding Rate of Methanol on Cell Growth 130
4.4.2.2 Effect of Sorbitol and Feeding Rate of Methanol on Recombinant Protein Production 132
4.4.2.3 Effect of Methanol Feeding Rate on Sorbitol Consumption Rate 133
4.4.2.4 Effect of Sorbitol and Feeding Rate of Methanol on Total Protease Production 134
4.4.2.5 Effect of Sorbitol and Feeding Rate of Methanol on Alcohol Oxidase Production 135
4.4.2.6 Effect of Sorbitol and Feeding Rate of Methanol on Amino and Organic Acid Concentration Profiles 136
4.4.2.7 Oxygen Transfer Characteristics of the Bioprocess 140
4.4.2.8 Yield Coefficients and Specific Rates of the Bioprocess 143
4.5 Stoichiometric Modeling of the Bioprocess and Metabolic Flux Analysis 146
4.5.1 Metabolic Flux Analysis for M-0.03 Condition 158
4.5.2 Metabolic Flux Analysis for MS-0.02 Condition 159
4.5.3 Metabolic Flux Analysis for MS-0.03 Condition 159
4.5.4 Metabolic Flux Analysis for MS-0.04 Condition 160
4.5.5 Metabolic Flux Distribution at Specific Nodes 161
4.6 Kinetic Modeling of the Bioprocess 163
4.6.1 Defining the Structure of the Model 163
4.6.2 Material Balances and Rate Equations 165
4.6.3 Finding the Kinetic Constants 173
4.6.4 Validation of the Model 184
5. CONCLUSION 189
APPENDICES 222
A. CHEMICALS AND SUPPLIERS 223

B. LABORATORY EQUIPMENT AND SUPPLIERS 226
C. BUFFERS AND STOCK SOLUTIONS 228
D. DNA SEQUENCES AND PLASMIDS 233
E. GROWTH MEDIA 239
F. MOLECULAR WEIGHT MARKERS 244
G. AMINO ACID CODONS, ABBREVIATIONS AND CODON USAGE 245
H. PROPERTIES OF DESIGNED PRIMERS 247
I. PROTOCOL FOR SOUTHERN BLOTTING 249
J. PROTOCOL FOR HIS-TAG PURIFICATION 252
K. CALIBRATION OF PROTEIN CONCENTRATION 254
L. SDS-PAGE AND WESTERN BLOTTING PROTOCOLS 255
M. STAINING PROTOCOL FOR SDS-POLYACRYLAMIDE GELS 258
N. CALIBRATION OF METHANOL, GLYCEROL, SORBITOL CONCENTRATIONS 259
O. METABOLIC REACTIONS FOR *P. pastoris* 261

CURRICULUM VITAE 268

LIST OF TABLES

TABLES

1.1 List of some fermentation products and market values in the year 2000. .. 2
2.1 Advantages and disadvantages of *P. pastoris* 10
3.1 Strains and plasmids used in this study.....56
3.2 Primers used in this study and their sequences 58
4.1 Transesterification reaction conditions for biodiesel synthesis, using canola oil (C) and sunflower oil (S).....108
4.2 Overall bioprocess yields attained for the rHuEPO production.....115
4.3 Experimental design for investigating the combined effect of salts on cellular growth. 117
4.4 The experimental plan for bioreactor experiments 129
4.5 Parameters used in equation 4.1 129
4.6 Variation in amino acid concentration profiles with cultivation time, for runs with different feeding profiles 138
4.7 Variation in organic acid concentration profiles with cultivation time, for runs with different feeding profiles 139
4.8 The variations in oxygen transfer parameters with different methanol feeding profiles 142
4.9 Overall yield coefficients 144
4.10 Variation in specific rates throughout the bioprocesses.....144
4.11 Metabolic flux distributions of rHuEPO production by *P. pastoris* at different feeding conditions 150
4.12 ATP generated throughout the bioprocess.....157
4.13 Summary of the proposed structured model equations for fed-batch production of rHuEPO by *P. pastoris* 172
4.14 Initial conditions and fixed parameters used in the proposed structured model 173
4.15 The kinetic constants of the specific AOX formation rate equation for different feeding conditions. 177
4.16 The kinetic constants of the specific protease formation rate equation for different feeding conditions. 178

4.17 The kinetic constants of the specific protease formation rate equation for different feeding conditions. 183
4.18 Summary of the final structured model equations for fed-batch production of rHuEPO by *P. pastoris* 185
4.19 The kinetic constants of the model 186

E.1 Composition of the various macronutrient solutions 241
E.2 Composition of the various trace element solutions 242
G.1 Amino Acid Codons and Abbreviations.....245
G.2 Codon usage by *P. pastoris.* 246
H.1 Thermodynamic properties of the designed primers. 247

LIST OF FIGURES

FIGURES

2.1 Primary structure of human erythropoietin.. 8
2.2 Tertiary structure of human erythropoietin. 8
2.3 Overview of yeast and mammalian N-linked glycosylation pathway. 16
2.4 The basic series of steps for cloning a segment of foreign DNA to obtain a recombinant DNA molecule, the recombinant plasmid. 18
2.5 The Polymerase Chain Reaction 20
2.6 Schematic representation of MALDI-ToF MS working principle in reflectron mode. 34
2.7 Variation of dissolved oxygen concentration with time in dynamic measurement of K_{La}.....43
2.8 Evaluting K_{La} using the Dynamic Method. 43
2.9 Simplistic overview of yeast carbon metabolism 49
3.1 Scale up steps and the pilot scale bioreactor system. 67
4.1 Flowchart of the research plan, for the development of the recombinant *Pichia pastoris* producing extracellular rHuEPO. 78
4.2 Agarose gel electrophoresis of pPICZαA extracted from *E. coli* Top10.....79
4.3 Agarose gel electrophoresis results A. Plasmid (pENTR221::EPO) extracted from *E. coli* IOH44362. B. EPO gene amplified by PCR, using pENTR221::EPO as template and the primer set EPO-F3-1 and EPO-R3....80
4.4 Schematic representation of epo amplification, integration of specific recognition sites by two-step-PCR, and construction of pPICZαA::*epo* plasmid. 81
4.5 Agarose gel electrophoresis of products of digestion and ligation reactions. 83
4.6 Agarose gel electrophoresis of double digestion by *Eco*RI and *Xba*I of the plasmids from the twelve selected *E. coli* transformants. 84
4.7 Schematic representation of pPICZαA::epo integration into *P. pastoris* genome. 85
4.8 Agarose gel electrophoresis of PCR amplification of EPO insert from genomic DNA of *P. pastoris* transformants. 86

4.9 Agarose gel electrophoresis view of the *Bam*HI digested genomic DNA of *P. pastoris* transformants, to be used in Southern blotting.. 88
4.10 Southern blot hybridization of digested *P. pastoris* transformants' genomic DNA, using epo insert as the probe. 88
4.11 Western blot analysis of the proteins produced by *P. pastoris* transformants. 90
4.12 Optical density at 600 nm of the BMMY medium during production phase of *P. pastoris* cultivation. 91
4.13 SDS-PAGE gel of *P. pastoris* production medium. 92
4.14 EPO amino acid sequence. 93
4.15 SDS-PAGE gel view of the proteins bound to the IMAC resins but eluted by step-elution prior to rHuEPO elution **.....** 95
4.16 SDS-PAGE gel view before and after optimization studies.. 97
4.17 (a) Silver stained SDS-PAGE view of rHuEPO produced by *P. pastoris* E17. (b) Western blot view of rHuEPO produced by *P. pastoris* E17. 98
4.18 A Target plate and magnified photographs of the sample spots for MALDI-ToF MS.. 100
4.19 MALDI mass spectra of glycans of rHuEPO (released by N-glycanase treatment) on an S-DHB matrix, detected in $[M+Na]^+$ form. 103
4.20 MALDI-ToF mass spectrum of the deglycosylated rHuEPO (m/z=20376) on sinapinic acid matrix. 104
4.21 MALDI-ToF mass spectrum of the native rHuEPO, deglycosylated and factor Xa protease digested rHuEPO, (m/z=9273) on sinapinic acid matrix, detected in $[M+2H]^{2+}$ form. 106
4.22 Variations in the glycerol concentration with the cultivation time in media containing different sources of glycerol 111
4.23 Variations in the cell concentration with the cultivation time in media containing different sources of glycerol 111
4.24 A. Dot blot analysis at t=54h of the bioprocess. B. RHuEPO concentrations measured by HPCE in different media **.....**113
4.25 Effects of macro and micronutrients on cell concentration. 118
4.26 Variation in cell concentration in BSM+PTM1 media, induced with 1% methanol at t=0 h only, and containing different initial sorbitol concentrations. 119

4.27 Dot blot analysis view of 10 µl supernatant samples taken at different cultivation times of the bioprocess, from media containing different initial sorbitol concentrations.....120
4.28 Variation in cell concentration with cultivation time in BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol and different methanol induction strategy...121
4.29 Variation in cell concentration with cultivation time in BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol and different methanol induction strategy...121
4.30 The variation in cell concentration with cultivation time in media with different initial sorbitol concentrations 123
4.31 Variation in cell concentration with initial sorbitol concentration, at t=48 h 123
4.32 Variation in cell concentration with respect to medium pH. 124
4.33 Silver stained SDS-PAGE gel view of extracellular proteins produced by *Pichia pastoris* in shake-bioreactor experiment, to observe the effect of medium pH on rHuEPO production.....125
4.34 A screenshot of the dissolved oxygen levels during two bioreactor operations.....127
4.35 The predetermined feed profiles for methanol, calculated for each specific growth rate (μ_o) aimed. 130
4.36 a. Variation in cell concentration with cultivation time in the precultivation phases b. Variation in cell concentration with cultivation time in the production phase of bioprocesses 131
4.37 SDS-PAGE gel view of extracellular production medium proteins from bioprocesses with different feeding profiles.....132
4.38 Variation in recombinant protein concentration with cultivation time in the production phase of bioprocesses with different feeding profiles.....133
4.39 Variation in sorbitol concentration with cultivation time for different feeding profiles 134
4.40. Variation in total protease concentration with cultivation time for different feeding profiles. 135
4.41 Variation in alcohol oxidase activity with cultivation time for different feeding profiles 136
4.42 Variation in lactic acid concentration with cultivation time for different feeding profiles. 137
4.43 The metabolic pathway map of *P. pastoris* producing rHuEPO.....147

4.44 Variation in biomass and metabolite concentrations with the cultivation time for MS-0.03 condition, and the three periods of the bioprocess. 149
4.45 Normalized metabolic flux distributions at Pyr (at t=15 h) and G3P (at t=9 h) nodes for the bioprocesses with different with feeding conditions. 162
4.46 Diagram of the proposed model, showing the compartments 164
4.47 The simplified problem scheme for modeling of the fed-batch bioprocess. 165
4.48 Determination of ms by fitting the model to experimental data for MS-0.03 condition. 174
4.49 The graphs showing the rate equation correlation for specific AOX formation rate (q_{AOX}) for each fed-batch run. 176
4.50 The graphs showing the linear relationship between the kinetic constants and the design parameter μ_0. 177
4.51 The correlation graph showing the linear relationship between the kinetic constant, k_{p2}, and the design parameter μ_0.....178
4.52 Comparison of the experimental data and modeling results of the variation in extracellular protease concentration with cultivation time. 180
4.53 The graphs showing the rate equation correlation for specific rHuEPO formation rate (q_{rp}) for all conditions. 181
4.54 Comparison of the experimental data and modeling results of the variation in rHuEPO concentration with cultivation time for MS-0.02 condition, when degradation by proteases is neglected from the model. 181
4.55 The graphs showing the rate equation correlation for specific degradation rate (q_d) for MS-0.04 condition. 182
4.56 The correlation graph showing the linear relationship between the kinetic constants, k_{d1}, k_{d2} and the design parameter μ_o. 183
4.57 Model prediction of experimental data. 187
4.58 Comparison of the experimental data and modeling results of the variation in rHuEPO concentration with cultivation time. 188

D.1 Schematic representation of pPICZαA vector as supplied from Invitrogen. 234
D.2 Schematic representation of pENTR 221 as supplied from Invitrogen 238
F.1 A. DNA Ladder, Hyperladder I (BioLine) ; B. Precision Plus Dual Color Prestained Protein MW Marker (BioRad) C. PageRuler™ Prestained Protein Ladder (Fermentas) 244

K.1 Standard curve for Bradford Assay 254
N.1 Calibration curve obtained for methanol concentration; analysis was performed by HPLC. 259
N.2 Calibration curve obtained for glycerol concentration; analysis was performed by HPLC. 259
N.3 Calibration curve obtained for sorbitol concentration; analysis was performed by HPLC. 260

NOMENCLATURE

C	Concentration in the medium	g L^{-1} or mol m^{-3}
C_O^*	Saturated dissolved oxygen concentration	mol m^{-3}
Da	Damköhler number (=OD / OTR_{max} ; Maximum possible oxygen utilization rate per maximum mass transfer rate)	
DO	Dissolved oxygen	%
E	Enhancement factor (=K_La / K_La_o); mass transfer coefficient with chemical reaction per physical mass transfer coefficient	
k	Kinetic rate constant	
K_La	Overall liquid phase mass transfer coefficient	s^{-1}
K_La_0	Physical overall liquid phase mass transfer coef.	s^{-1}
Ks	Saturation constant	g L^{-1}
m_s	Maintenance coefficient	g g^{-1} h^{-1}
N	Agitation rate	min^{-1}
OUR	Oxygen uptake rate	mol m^{-3} sec^{-1}
OTR	Oxygen transfer rate	mol m^{-3} sec^{-1}
OD	Oxygen demand	mol m^{-3} sec^{-1}
Q	Feed inlet rate	L h^{-1}
q	Specific formation or consumption rate	g g^{-1} h^{-1}
r	Formation or consumption rate	g L^{-1} h^{-1}
t	Cultivation time	h
T	Bioreaction medium temperature,	°C
U	One unit of an enzyme	
V	Volume of the bioreactor	L
Y	Yield (overall)	g g^{-1}

Greek Letters

ρ	Density	g L^{-1}
η	Effectiveness factor (=OUR/OD)	
μ_0	Desired specific growth rate	h^{-1}

$\mu_{s,max}$	Maximum specific growth rate on sorbitol	h^{-1}
μ_t	Total specific growth rate	h^{-1}
λ	Wavelength	nm

Subscripts

0	Refers to initial condition
AOX	Refers to alcohol oxidase
c	Refers to cell
d	Refers to degradation
G	Refers to glycerol
M	Refers to methanol
O	Refers to oxygen
p	Refers to protein
pro	Refers to protease
R	Refers to bioreaction medium
rp	Refers to recombinant protein
S	Refers to sorbitol or substrate
X	Refers to cell

Abbreviations

AOX	Alcohol oxidase
CDW	Cell dry weight
CHO	Chineese Hamster ovary
DNA	Deoxyribonucleic acid
EPO	Human erythropoietin hormone
HPCE	High pressure capillary electrophoresis
HPLC	High pressure liquid chromatography
m/z	Mass/ionic charge
MALDI-Tof	Matrix-assisted laser desorption/ionization time-of-flight
MFA	Metabolic flux analysis
MS	Mass spectrometry
PCR	Polymerase chain reaction
pI	Isolelectric point
PPP	Pentose phosphate pathway
Q-ToF	Quadrupole time-of-flight
R#	Reaction number

rHuEPO	Recombinant human erythropoietin
SDS-PAGE	Sodium dodecylsulfate-polyacrylamide gel electrophoresis
TCA	Tricarboxylic acid
Z	Objective function

Abbreviations Used in Metabolic Flux Anlaysis

Ac	Acetate
AcCoA	Acetyl coenzyme A
AcCoAm	Acetyl coenzyme A, mitochondrial
Acet	Acetaldehyde
ADP	Adenosine 5'-diphosphate
aKG	α-ketoglutarate, mitochondrial
Ala	L-Alanine
AMP	Adenosine 5'-monophosphate
Arg	L-Arginine
Asn	L-Asparagine
Asp	L-Aspartate
ATP	Adenosine 5'-triphosphate
CaP	Carbamoyl-phosphate
CARBH	Carbohydrates
CDP	Cytidine 5'-diphosphate
Chor	Chorismate
Cit	Citrate, mitochondiral
CMP	Cytidine 5'-monophosphate
CO_2	Carbondioxide
CTP	Cytidine 5'-triphosphate
Cys	L-Cysteine
dATP	2'-Deoxy-ATP
dCTP	2'-Deoxy-CTP
dGTP	2'-Deoxy-GTP
DHF	Dihydrofolate
dTTP	2'-Deoxy-TTP
E4P	Erythrose 4-phosphate
EPO	Erythropoietin
EtOH	Ethanol
F10THF	N^{10}-Formyl-THF

F6P	Fructose 6-phosphate
FA	Fatty acids
$FADH_2$	Flavine adenine dinucleotide, reduced
For	Formate
Form	Formaldehyde
Fruc	D-Fructose
Fum	Fumarate, mitochondrial
G3P	Glyceraldehyde 3-phosphate
G6P	Glucose 6-phosphate
GDP	Guanosine 5'-diphosphate
Gln	L-Glutamine
Glu	L-Glutamate
Gly	L-Glycine
GMP	Guanosine 5'-monophosphate
GTP	Guanosine 5'-triphosphate
H_2S	Hydrogen sulfide
His	L-Histidine
HSer	Homoserine
ICit	Isocitrate, mitochondrial
Ile	L-Isoleucine
IMP	Inosinemonophosphate
Kval	Ketovaline
Lac	Lactic acid
Leu	L-Leucine
Lys	L-Lysine
MaCoA	Malonyl-CoA
Mal	Malate, mitochondrial
MeOH	Methanol
Met	L-Methionine
MeTHF	N^5- N^{10}-methenyl-THF
MetTHF	N^5- N^{10}-methylene-THF
MTHF	N^5- methyl-THF
NADH	Nicotinamide-adeninedinucleotide (reduced)
NADHm	Nicotinamide-adeninedinucleotide, mitochondrial (reduced)
NADPH	Nicotinamide-adeninedinucleotide phosphate (reduced)
NADPHm	Nicotinamide-adeninedinucleotide phosphate, mitoch. (red.)

NH_4	Ammonia
OA	Oxaloacetate
OLE	Oleic acid
PA	Phosphatidate
PC	Phosphatidylcholine
PE	Phosphatidylethanolamine
PEP	Phosphoenolpyruvate
3PG	3-Phospho-D-glycerate
Phe	L-Phenylalanine
Pi	Inorganic ortophosphate
PINS	1-Phosphatidyl-D-myo-inositol
PLLM	Palmitoleic acid
PLM	Palmitic acid
PPi	Inorganic pyrophosphate
PRAIC	1-(5'-Phosphoribosyl)-5-amino-4-imidazolecarboxamide
Pro	L-Proline
PRPP	5-Phospho-D-ribosylpyrophosphate
PS	Phosphatidylserine
Pyr	Pyruvate
R5P	Ribulose 5-phosphate
Rib5P	Ribose 5-phosphate
S7P	Sedoheptulose-7-phosphate
Ser	L-Serine
Sorb	Sorbitol
STE	Stearate
Suc	Succinate, mitochondrial
SucCoA	Succinate coenzyme A, mitochondrial
THF	Tetrahydrofolate
Thr	L-Threonine
Trp	L-Tyrptophan
Tyr	L-Tyrosine
UDP	Uridine 5’-diphosphate
UMP	Uridine 5’-monophosphate
UTP	Uridine 5’-triphosphate
Val	L-Valine
Xyl5P	Xylulose 5-phosphate

CHAPTER 1

INTRODUCTION

Industrial biotechnology is a rapidly growing field, with the interaction of biological sciences and engineering disciplines. With the increasing shift towards a bio-based economy, there is rising demand for developing efficient cell-factories that can produce fuels, chemicals, pharmaceuticals, and even food ingredients (Nielsen and Jewett, 2008). The exploitation of cells as micro-reactors within bioreactors (Çalık et al., 1999) is generally referred to as a fermentation process. The spectrum of fermentation products (Table 1.1) have increased even more in the last few decades with the advances in recombinant DNA technology, as it is now possible to produce bio-molecules in an organism other than its natural source (Nielsen et al., 2003). The most remarkable change happened in the recombinant therapeutic proteins market, which reached to a value about $32 billion in 2003, as recombinant DNA technology directly facilitated the production of significant quantities of the biologically active therapeutic proteins.

Recombinant human erythropoietin (rHuEPO), is one of the recombinant therapeutic proteins approved by the U.S. Food and Drug Administration (FDA), and it is now widely used for the treatment of anemia associated with renal failure, cancer, prematurity, chronic inflammatory disease and human immunodeficiency virus infection (Jelkman, 1992), with global sales in excess of $3.6 billion in 2000. Erythropoietin (EPO) is a glycoprotein human hormone that chiefly regulates the production of red blood cells, and is produced primarily by the kidney in the adult and by the liver during fetal life (Jacobson et al., 1957;

Table 1.1 List of some fermentation products and market values in year 2000 (Adapted from Nielsen et al., 2003).

Product	Typical organism	Market value
Recombinant Proteins		
Erythropoietin	Chinese Hamster Ovary cells	3.6 billion US$
Insulin	*Saccharomyces cerevisiae*	3 billion US$
Human growth hormone	*Escherichia coli*	1 billion US$
Interferons	*Escherichia coli*	2 billion US$
Primary Metabolites		
Ethanol	*Saccharomyces cerevisiae*	12 billion US$
Lactic acid	*Zymomonas mobilis*	0.2 billion US$
Citric acid	*Rhizopus oryzae*	1.5 billion US$
Glutamate	*Corynebacterium glutamicum*	1 billion US$
Lysine	*C. glutamicum*	0.5 billion US$
Phenylalanine	*C. glutamicum*	0.2 billion US$
Secondary Metabolites		
Penicillins	*Penicillium chrysogenum*	4 billion US$
Taxol	Plant cells	1 billion US$
Whole Cells		
Single cell protein	Methylotrophic bacteria	
Baker's yeast	*Saccharomyces cerevisiae*	
Enzymes		
Detergent enzymes	*Bacilli, Aspergilli*	0.6 billion US$
Starch industry enzymes	*Bacilli, Aspergilli*	0.2 billion US$
Polymers		
Xanthan gum	*Xanthomonas campestris*	0.4 billion US$
DNA		
Gene therapy	*Escherichia coli*	

Zanjani et al., 1977). Decreased production of EPO, due to kidney failure, results in anemia.

The large quantities of the hormone required to satisfy clinical demand are currently met by recombinant expression in mammalian cells, namely Chinese hamster ovary (CHO) cells (Egrie, 1990). However, there are significant disadvantages to the use of mammalian cell cultures with respect to efficiency and cost (Fernandez and Hoeffler, 1999), and there is a strong need to find alternative systems for a more efficient production of EPO.

Production of EPO in bacterial hosts has been reported for *Escherichia coli* (Leehaung, 1984; Bill et al., 1995) and for *Bacillus brevis* (Nagao et al., 1997), and in a eukaryote host - the baker's yeast, *Saccharomyces cerevisiae* (Elliot et al., 1989). *Drosophila melanogaster* Schneider 2 cells were also recently employed (Kim et al., 2005) in an attempt to achieve a simpler eukaryotic expression system, as an alternative to CHO cells, for production of rHuEPO; the *N*-glycan structures of rHuEPO so produced were investigated. However, prokaryotes do not possess the ability to glycosylate proteins, while even microbial eukaryotes can affect that process; although, admittedly, *S. cerevisiae* hyperglycosylates proteins (Walker, 1998). Thus, currently none of these systems seem eligible to take the place of CHO cells.

Pichia pastoris has become popular as a useful alternative to *S. cerevisiae*, which is the most commonly, used and studied yeast for production of heterologous proteins. There are several reasons that account for the popularity of the *P. pastoris* expression system: its ability to produce foreign proteins at high levels, either extracellularly or intracellularly; its facility in performing many post-translational modifications (e.g. glycosylation without the hyperglycosylation of *S. cerevisiae*), correct disulfide bond formation, and proteolytic processing); the availability of the alcohol oxidase I (*AOX1*) promoter (known to be one of the strongest and most tightly regulated eukaryotic promoters) for controlled gene expression; the ability to stably integrate expression plasmids at specific sites in the *P. pastoris* genome in either single or multiple copies; and its ability to grow to a very high cell density in bioreactors (Cereghino and Cregg 1999; 2000). The engineering of *P. pastoris* has been reported recently, to produce humanized glycoproteins and it has been shown that a recombinant rat EPO secreted by this

particular host was able to raise hematocrit levels in mice where as the wild type *P. pastoris* had much less activity (Hamilton et al. 2006).

In this PhD thesis study, it was aimed to develop a bioprocess using the *Pichia pastoris* expression system as an alternative to mammalian system used in the industry, for the production of the therapeutically important glycoprotein, erythropoietin (EPO). In this context, the EPO gene from the human cell-line was recombined to a *P. pastoris* strain for the first time. The recombinant microorganism producing rHuEPO, with native *N*- and *C*-termini, was developed and the recombinant product produced by this microorganism was purified and structurally analyzed. Thereafter, a defined fermentation medium was designed, and in addition to the medium designed, an an alternative and cheap carbon source to be used in *P. pastoris* fermentation processes, the byproduct of biodiesel industry, crude glycerol without prior purification, was suggested for the first time in this study. Furthermore, the optimum bioprocess design parameters, i.e. medium components and pH, were determined for increased production efficiency, and the carbon source feeding strategy was investigated in pilot scale bioreactors. The data collected in pilot scale bioreactor experiments were then used in the stoichiometric and kinetic modeling, in order to determine the intracellular reaction rates and consequently to identify the potential metabolic bottlenecks in the synthesis of rHuEPO, as well as to better describe and predict the process outcomes.

CHAPTER 2

LITERATURE SURVEY

2.1 Bioprocess Development

Since the ancient times where microorganisms were used to transform biological materials for production of fermented foods, bioprocesses have been developed for an enormous range of commercial products, from relatively cheap materials such as industrial alcohol and organic solvents, to expensive specialty chemicals such as antibiotics, therapeutic proteins and vaccines. Reflecting the complex nature of biological systems, development of a complete industrial bioprocess requires many stages (Doran, 1995). Once the product and process goals have been determined, the first stages of bioprocess development are concerned with selecting a proper host cell and performing genetic manipulations if necessary. After cloning, the growth and production characteristics of the cells must be obtained as a function of culturing conditions. Thus, medium composition, pH and temperature allowing optimal growth and productivity are determined in small scale experiments. During these initial production experiments, purification and structural analysis of the produced recombinant protein must be carried out simultaneously, in order to guide the subsequent stages. When the optimal culture conditions for production are known, scale-up of the bioprocess starts. The first stage is usually a 1- or 2-liter bench-top bioreactor equipped with instruments for measuring and adjusting temperature, pH, dissolved oxygen concentration, stirrer speed and other process variables. As cultures can be more closely monitored in bioreactors than in shake flasks, better control over the bioprocess is possible. At this stage, information is collected about the oxygen requirements of the cells, foaming characteristics and many other parameters,

and limitations imposed by the reactor on activity of the organism must be identified. The situation is assessed by calculation of the bioprocess characteristics such as mass-transfer coefficients, oxygen uptake and transfer rates, yields, productivity and many others. It must also be decided whether the culture is best operated as a batch, fed-batch or continuous process (Shuler and Kargı, 2002). Moreover, stoichiometric and kinetic models formed at this stage, will be useful in determining the bottlenecks of the process and better describe and predict the process outcomes. All these stages were considered in this bioprocess development study, whereas further stages of scale-up and product recovery were not included.

2.1.1 Product: Erythropoietin (EPO)

Erythropoietin (EPO), being a glycoprotein, is a human hormone that control red blood cell production through the promotion of survival, proliferation and differentiation of the erythroid progenitors in the bone marrow. It is produced primarily by the kidney in the adult and by the liver during fetal life (Jacobson et al., 1957; Zanjani et al., 1977). Erythropoietin hormone was first purified in small amounts from human urine in 1977 (Miyake et al., 1977). Later in 1985, through the use of recombinant DNA technology, human EPO gene was cloned, sequenced and expressed in biologically active form in mammalian cells (Chinese Hamster Ovary (CHO) cells and Baby Hamster Kidney (BHK) cells) (Jacobs et al., 1985; Lin et al., 1985). Decreased production of EPO, due to kidney failure, results in anemia. Recombinant human erythropoietin (rHuEPO), is one of the recombinant therapeutic proteins approved by the U.S. Food and Drug Administration (FDA), and it is now widely used for the treatment of anemia associated with renal failure, cancer, prematurity, chronic inflammatory disease and human immunodeficiency virus infection (Jelkmann, 1992), with global sales in excess of $3.6 billion in 2000.

Molecular cloning studies have indicated that there is a single copy of the erythropoietin gene on human chromosome 7 in q11-22 region (Law et al., 1986). The gene spans a 5.4kb region on human chromosome with five exons and four introns (Maiese et al., 2005). The physicochemical properties and the carbohydrate structure of EPO and rHuEPO were characterized in detail (Lai et al., 1986; Recny et al., 1987; Sasaki et al., 1987; Sasaki et al., 1988; Takeuchi et al., 1988; Tsuda et al., 1988). It was reported that, the human EPO gene

encodes a protein of 193 amino acids, however the mature hormone is composed of 165 amino acids (Figue 2.1). The first 27 amino acids are presumed to be the hydrophobic leader sequence that is later cleaved during secretion of the protein (Lin et al., 1985; Egrie, 1990) and the amino acid at position 166 is cleaved off later, but again prior to secretion. EPO in its structure, has two disulfide bonds (between 7 and 161, and between 29 and 33), and four oligosaccharide chains (three *N*-linked at Asn 24, Asn 38, Asn 83, and one *O*-linked at Ser126), which constitutes about 40% of the molecular weight of EPO. The molecular weight of glycosylated EPO was determined to be 34-38.5 kDa, whereas that of the peptide chain was about 18 kDa. The tertiary structure of native EPO without glycosylation is shown in Figure 2.2. The role of the carbohydrate moieties were widely studied (Takeuchi et al., 1989; Takeuchi et al., 1990; Tsuda et al., 1990; Higuchi et al., 1992; Skibeli et al., 2001; Elliott et al., 2004), and many methods using capillary electrophoresis, chromatography and mass spectrometry were developed for analyzing the oligosaccharide profiles of EPO (Lopez-Soto-Yarritu et al., 2002; Guile et al., 1996; Stübiger et al., 2005). In general, the carbohydrate chains of a glycoprotein affect its conformation leading to changes in intracellular transport and secretion, solubility, susceptibility to proteases, *in vivo* half-lie by altering routes of clearance, and receptor binding properties.

2.1.2 Microorganism Selection

The first crucial step in the development of a bioprocess with efficient protein expression is the choice of the host microorganism. Potential hosts should give sufficient yields, be able to secrete large amounts of protein, be suitable for industrial fermentations, produce a large cell mass per volume quickly and on cheap media, be considered safe based on historical experience or evaluation by regulatory authorities, and should not produce harmful substances or any other undesirable products (Kirk and Othmer, 1994). When recombinant production processes began to be developed about 25 years ago, *E. coli* and *S. cerevisiae* were the most advanced and therefore mostly applied organisms, and now are established in industry and accepted by regulatory bodies. However, *E. coli* does not possess the ability to perform many post-translational modifications such as glycosylation, while *S. cerevisiae* hyper-glycosylates proteins (Walker, 1998) which reduces secretion rates and potentially alters protein functionality.

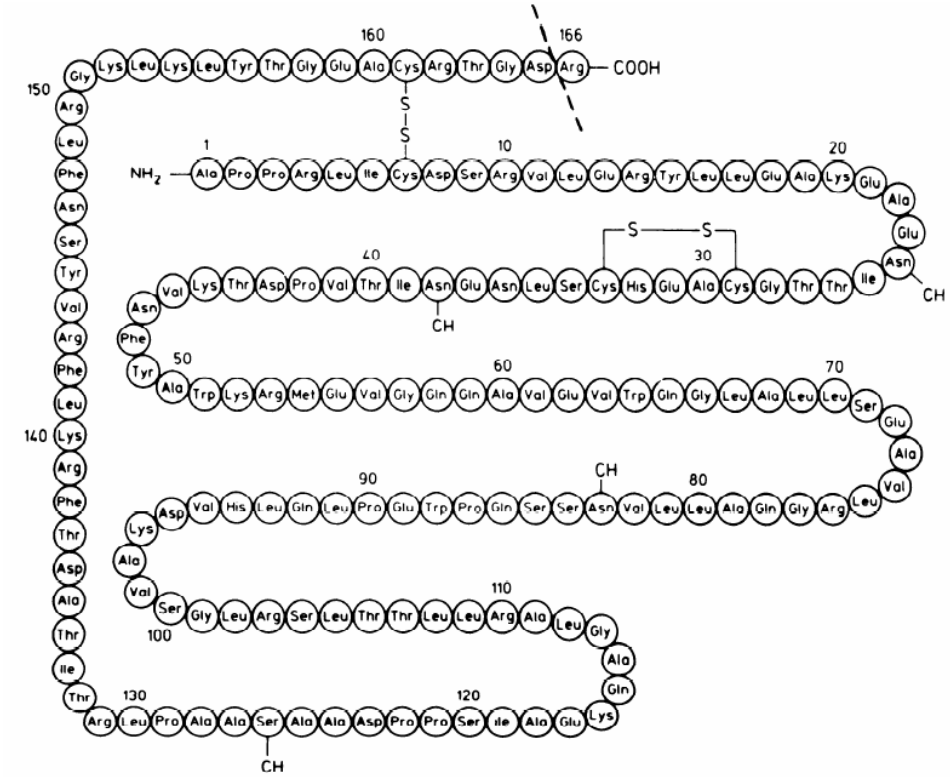


Figure 2.1 Primary structure of human erythropoietin. Mature hormone is composed of 165 amino acids, having lost carboxyterminal arginine residue by posttranslational modification. There are three *N*-linked glycosylation sites at aspartyl residues 24,38, and 83, and one *O*-linked glycosylation site at serine residue 126 (Jelkmann, 1992).

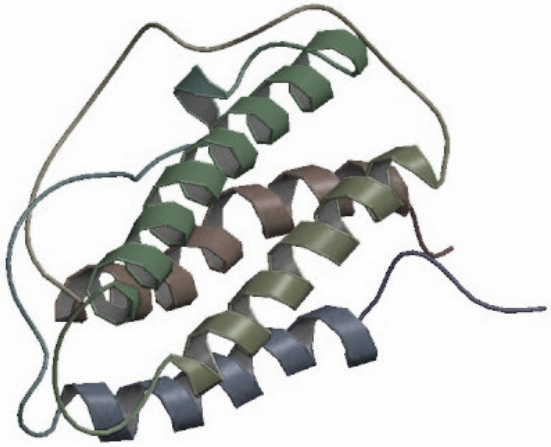


Figure 2.2 Tertiary structure of human erythropoietin. (http://www.rcsb.org /pdb/explore/explore.do?structureId=1BUY, Last accessed: September 2008).

Nevertheless, *Pichia pastoris* serves to be advantageous in these and many other aspects, and is at present the most frequently used species for heterologous protein expression in general.

2.1.2.1 *Pichia pastoris*

2.1.2.1.1 General Characteristics

Pichia pastoris, first employed by the Phillips Petroleum Company in the 1970s, is now considered to be superior to any other known yeast species with respect to its secretion efficiency and permits the production of recombinant proteins without intense process development (Schmidt, 2004). More than 500 proteins have been cloned and expressed using this system. Furthermore, it has been selected by several protein production platforms for structural genomics programs (Cos et al., 2006).

P. pastoris is a mesophilic microorganism from the yeast species. It is taxonomically classified under the Kingdom *Fungi*, Division *Eumycota*, Subdivision *Ascomycotina*, Class *Hemoascomycetes*, Order *Endomycetales*, Family *Saccharomycetaceae* and Genus *Pichia*. Yeasts are unicellular fungi which usually appear as oval cells 1-5 μm wide by 5-30 μm long, they have typical eukaryotic cell structure and generally have a thick polysaccharide cell wall, and they are facultative anaerobes. *Pichia pastoris* reproduce asexually, i.e., it is a homothallic ascomycetous yeast and remains haploid unless forced to mate (Cregg, 1999).

P. pastoris species has the advantages of eukaryotic cells and the simplicity of unicellular systems. The detailed advantages/disadvantages of *P. pastoris* are shown on Table 2.1 (Cregg, 1999; Daly and Hearn, 2005; Macauley-Patrick et al., 2005).

2.1.2.1.2 Glycerol, Methanol and Sorbitol Metabolisms

In standard processes with recombinant *P. pastoris* strains for the production of foreign proteins, *P. pastoris* is usually grown first on glycerol as carbon source, since biomass yield and maximum specific growth rate are higher for growth on glycerol than for growth on methanol. Moreover, during growth on glycerol, recombinant protein expression is repressed. Use of glucose as carbon source is

Table 2.1 Advantages and disadvantages of *P. pastoris*

Advantages	***Disadvantages***
• High yield and productivity • Strong promoter (AOX1) • Chemically defined media-simple, inexpensive formulation • Product processing like mammalian cells • Stable production strains • Low purification cost • High levels of expression of intracellular and secreted proteins • Eukaryotic post-translational modifications • No endotoxin problem • Non-pathogenic • Broad pH range: 3- 7 • Ability of utilizing methanol • Preference for respiratory growth rather than fermentative, a major advantage relative *S. cerevisiae.* • Crabtree-negative • Hyper-glycosylation is not as much as in *S. cerevisiae*.	• Potential of proteolysis, non-native glycosylation. • Long time for cell cultivation compared to bacteria • Monitoring methanol during a process is difficult in order to induce *AOX*1 promoter. • Since methanol is a petrochemical substance, it may be unsuitable for use in the food industry and also storing of this in industrial scale is undesirable because it is a fire hazard.

usually avoided, because byproduct ethanol is observed in higher amounts than during growth on glycerol (Macauley-Patrick et al., 2005). Residual ethanol concentrations repress the alcohol oxidase promoter, even at levels of around 10-50 mg L^{-1} (Inan and Meagher, 2001-a).

The glycerol catabolic pathway includes phosphorylation by a glycerol kinase, which leads to glycerol 3-phosphate, followed by oxidation to dihydroxyacetone phosphate (DHAP) by a FAD-dependent glycerol-3-phosphate dehydrogenase located on the outer surface of the mitochondrial inner membrane. The dihydroxyacetone phosphate formed enters the glycolytic pathway. A few yeast species have an alternative pathway for dissimilating glycerol which involves a NAD-dependent glycerol dehydrogenase and a dihydroxyacetone kinase. However this pathway seems to be of less importance (Nevoigt and Stahl, 1997). The related pathway reactions are given in Table 2.2.

Some essential enzymes required for the methanol metabolism are found at high levels only when the cells are grown on methanol (Cereghino et al., 2000). The enzyme alcohol oxidase (AOX), can account for up to 35% of total cell protein, but is virtually absent in cells grown on glucose, glycerol or ethanol. AOX catalyses the first step in the dissimilation of methanol, which consists in the oxidation of methanol to formaldehyde (Form) and hydrogen peroxide using molecular oxygen. This reaction takes place within specialized organelles called peroxisomes (Veenhuis et al., 1983) in order to avoid the toxic effect of hydrogen peroxide produced during this reaction (Sreekrishna and Kropp, 1996). In cells grown on methanol, peroxisomes can account for over 80% of the cell volume.

A portion of formaldehyde generated by AOX leaves the peroxisome and enters the dissimilatory pathway, and is oxidized to formate and carbon dioxide by two subsequent dehydrogenase reactions. Glutathione-dependent formaldehyde dehydrogenase catalyzes the production of formate (For), from which carbon dioxide is produced by the action of formate dehydrogenase. Besides the generation of energy in the form of NADH, this pathway plays an important role in the detoxification of formaldehyde in methylotrophic yeasts (Lee et al., 2002). The remaining formaldehyde enters the assimilatory pathway and reacts in a transketolase reaction with xylulose-5-phosphate (Xu5P) to yield dihydroxyacetone (DHA) and glyceraldehyde-3-phosphate (G3P), catalyzed by

the peroxisomal dihydroxyacetone synthase. The formed C_3 compounds are further assimilated within the cytosol and will serve as building blocks for biomass synthesis. Dihydroxyacetone is phosphorylated by a dihydroxyacetone kinase (Lüers et al., 1998) to dihydroxyacetone phosphate (DHAP). The related pathway reactions are summarized in Table 2.2.

The sorbitol on the hand, enters the metabolism from fructose-6-phosphate branch point, where sorbitol is first oxidized to fructose by sorbitol dehydrogenase, and followed by phosphorylation via a sorbitol kinase. The related pathway reactions are summarized in Table 2.2.

Table 2.2 Glycerol and methanol metabolism. The reactions and the responsible enzymes

A. Glycerol Metabolism	
Glyc + ATP → Glyc-3-P	EC 2.7.1.30 Glycerol kinase
Glyc-3-P + FAD → DHAP	EC 1.1.99.5 Glycerol-3-phosphate dehydrogenase
DHAP ↔ G3P	EC 5.3.1.1 Triose-phosphate isomerase
B. Methanol Metabolism	
MeOH → Form	EC 1.1.3.13 Alcohol oxidase
i. Assimilatory Pathway of Formaldehyde	
Xyl5P + Form → G3P + DHA	EC 2.2.1.3 Formaldehyde transketolase
ATP + DHA → ADP + DHAP	EC 2.7.1.29 Glycerone kinase
DHAP ↔ G3P	EC 5.3.1.1 Triose-phosphate isomerase
ii. Disssimilatory Pathway of Formaldehyde	
Form → For + NADH	EC 1.2.1.46 Formaldehyde dehydrogenase
For → NADH + CO2	EC 1.2.1.2 Formate dehydrogenase
C. Sorbitol Metabolism	
Sorb → Fruc + NADH	EC 1.1.1.15 D-sorbitol dehydrogenase
Fruc +ATP → ADP + F6P	EC 2.7.1.4 D-fructokinase

2.1.2.1.3 Expression with Alternative Promoters

The most important reason why *P. pastoris* is one of the most attractive microorganisms for the production of the recombinant proteins at high levels is because heterologous proteins can be expressed under the control of the strong, tightly regulated, and methanol-induced alcohol oxidase promoter, *AOX1*. A promoter is a regulatory region of DNA generally located upstream of a gene that allows transcription of the gene. In the wild-type strain, *AOX1* promoter drives the expression of the enzyme alcohol oxidase 1. AOX1 gene was first isolated by Ellis et al. (1985). The genome of *P. pastoris* actually contains two alcohol oxidase genes, *aox1* and *aox2* (Koutz et al., 1989). AOX1 is responsible for the majority of alcohol oxidase activity in the cell (Cregg et al., 1989) and large amounts of this enzyme are synthesized because alcohol oxidase has a very low affinity for oxygen (Couderc and Baratti, 1980; Veenhuis et al., 1983). Indeed, in cultures on methanol as the carbon source, AOX generally accounts for over 30% of the total cell protein (Couderc and Baratti, 1980). This strong *AOX1* promoter can therefore be used to drive the expression of recombinant proteins to high levels, up to 12 g L^{-1} of recombinant protein (Cregg et al., 1993).

With regard to methanol-utilizing ability, three phenotypes of expression strains of *P. pastoris* are present (Stratton et al., 1998):

- Methanol utilization positive (Mut^+): presence of functional AOX1 and AOX2 genes; growth on methanol at the wild-type rate.
- Methanol utilization slow (Mut^S): the AOX1 gene is disrupted, only the AOX2 gene is functional; methanol metabolism is dependent on the transcriptionally weaker AOX2 gene.
- Methanol utilization negative (Mut^-): both AOX1 and AOX2 genes are disrupted; methanol cannot be metabolized at all, but methanol is necessary for induction of recombinant protein expression.

The maximum specific growth rate on methanol is 0.14 h^{-1}, 0.04 h^{-1} and 0.0 h^{-1} for Mut^+, Mut^S and Mut^- *P. pastoris* strains, respectively (Stratton et al., 1998).

The *P. pastoris* glyceraldehyde-3-phosphate dehydrogenase (*GAP*) (Waterham et al., 1997), glutathione-dependent formaldehyde dehydrogenase (*FLD*1) (Shen et

al., 1998), and the more recently developed isocitrate lyase *(ICL1*) (Menendez et al., 2003) and 3-phosphoglycerate kinase (*PGK*1) (de Almeida et al. 2005) promoters are alternatives to the *AOX1* promoter, because use of methanol as carbon source can be avoided with these promoters. Since methanol is a fire hazard and is toxic for the food industry, avoiding methanol would be beneficial. Nevertheless, more studies are needed for the optimal expression under these promoters.

The recombinant *Pichia pastoris* developed in this study was a Mut^+ strain and expression was controlled by the *AOX1* promoter.

2.1.2.1.4 Secretion of Proteins and Post-Translational Modifications

A major advantage of *P. pastoris* compared to bacterial expression systems is that the yeast has the potential of performing many eukaryotic post translational modifications, such as (Cereghino and Cregg, 2000):

- Processing of signal peptide
- Folding
- Disulfide bond formation
- Certain types of lipid addition
- *O*- and *N*-linked glycosylation

To produce a protein extracellularly, a specific signal peptide should be fused with the protein, in order to direct the protein into secretory pathway. Directing a foreign protein to the extracellular medium serves as a substantial first step in protein purification, because it eliminates the cell disruption step. Moreover, *Pichia pastoris* secretes low levels of endogenous proteins, thus the secreted foreign protein constitute the vast majority of total protein in the medium. The α-factor prepro peptide from *S. cerevisiae* is the secretion signal used with most success. This signal sequence consists of a 19 amino acid signal (pre) sequence followed by a 66-residue (pro) sequence. The processing of this signal involves three steps. The first is the removal of the pre signal by signal peptidase in the endoplasmic reticulum. Second, Kex2 endopeptidase cleaves between Arg-Lys of

the pro leader sequence. This is rapidly followed by cleavage of Glu-Ala repeats by the Ste13 protein (Cereghino and Cregg, 2000). There are other alternative signal peptides, *P. pastoris* acid phosphatase (PHO1) signal, or the recombinant protein's native signal (Daly and Hearn, 2005). However, the most suitable secretion signal may not be found without experimental studies, due to many possible interactions that could result between the signal peptide and the protein itself. The α-factor prepro signal, already present in the vector purchased, was used in this study for secretion of EPO to the extracellular medium.

Folding and disulphide bond formation have been identified in some cases as the rate-limiting step in production of foreign proteins from *Pichia* (Hohenblum et al., 2004). There are other cases where highly disulphide bonded proteins have been successfully produced in *P. pastoris* (White et al., 1994).

Glycosylation

Glycosylation is the most common post-translational modification to proteins preceding protein secretion. Approximately 0.5-1% of the translated proteins in eukaryotes are glycoproteins. There is considerable variability in the length and type of oligosaccharides present on glycoproteins, and this structural heterogeneity is dependent not only on the characteristics of the target protein but also the host expression system chosen (Daly and Hearn, 2005). The correct glycosylation patterns on native mammalian proteins ensure their biological activity.

Whereas, bacterial expression systems are not capable of glycosylating proteins, yeasts like *Pichia pastoris* are capable of performing both types glycosylation, which are *O*- and *N*-linked glycosylation. However, the stepwise process of glycosylation in yeast differs from the mammalian glycosylation (Figure 2.3). Yeast glycosylation machinery generally causes hyperglycosylation; that is, it adds *N*-linked high-mannose oligosaccharides to proteins. Nevertheless, relative to the oligosaccharide structures on *S. cerevisiae*-secreted proteins, several differences are seen in *P. pastoris*-produced proteins, the most important of which is the frequent absence of hyperglycosylation (Cereghino and Cregg, 2000). In terms of *O*-linked oligosaccharides, while yeasts add mannose residues only, mammalian cells have the choice of N-acetylglucoseamine, galactose and

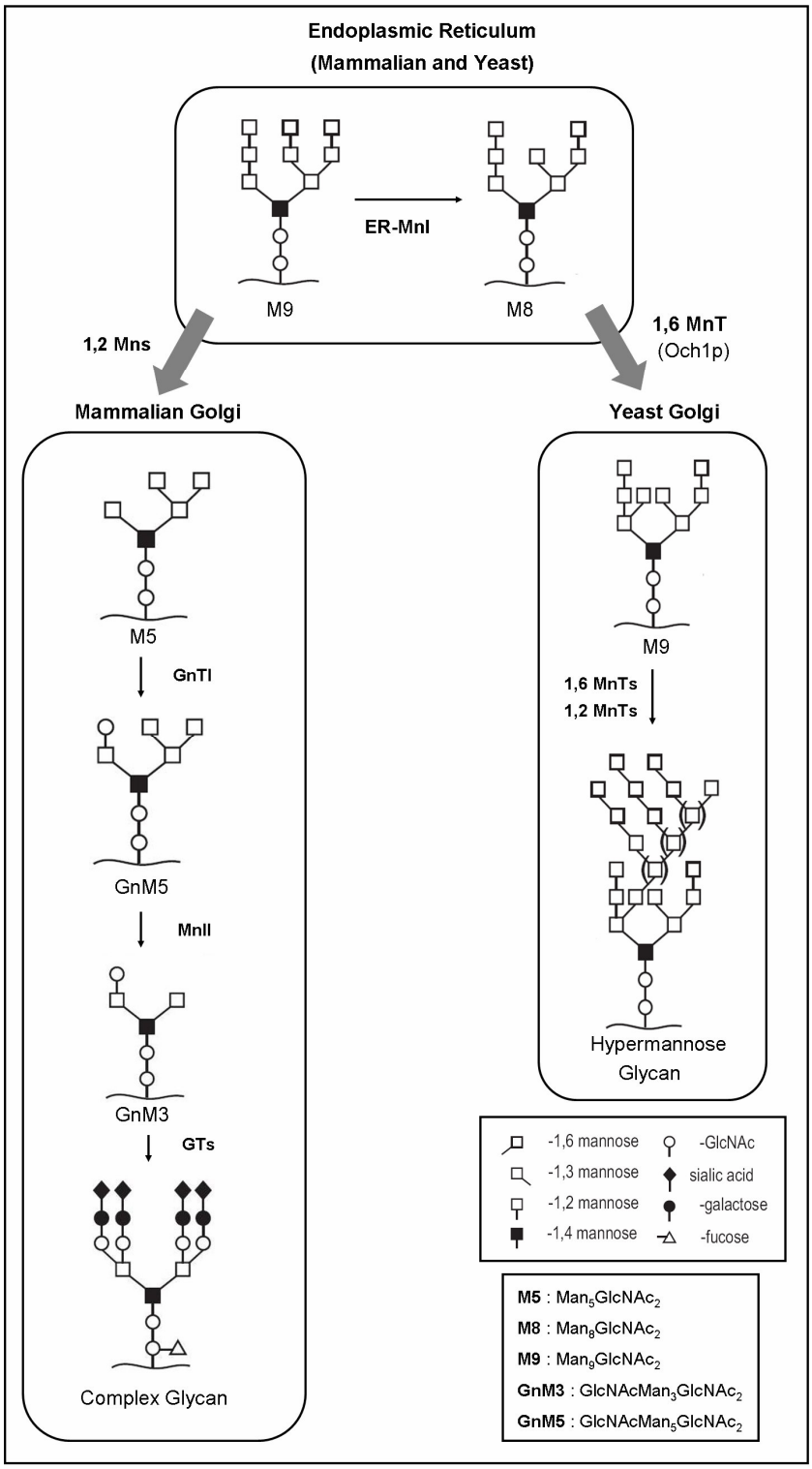


Figure 2.3 Overview of yeast and mammalian *N*-linked glycosylation pathway. ER-MnI: Endoplasmic reticulum mannosidase I; 1,2-Mns: 1,2-mannosidase; MnII: mannosidase II; GnTI: 1,2-N-acetylglucosaminyltransferase I; 1,2-MnT: 1,2-mannosyltransferase; 1,6-MnT: 1,6-mannosyltransferase; GTs: glycosyl transferases (various); Man: Mannose; GlcNAc: N-acetylglucosamine. (Adapted from Hamilton et al., 2006 and Vervecken et al., 2004).

sialic acid. Unlike *N*-glycosylation, which is known to be crucial for protein function, relatively little is known about *O*-glycosylation and its biological role.

Several approaches to humanizing yeast *N*-glycosylation pathways of *Pichia pastoris* have been attempted over the past decade (Callewaert et al., 2001; Choi et al., 2003; Bobrowicz et al., 2004) and advances in the glycoengineering of yeast and the expression of therapeutic glycoproteins with humanized *N*-glycosylation structures have shown significant promise (Hamilton et al., 2006). Thus, once studies and trials on engineering of the synthetic glycosylation pathway of *Pichia pastoris* is complete, the therapeutic protein production with human-type glycosylation will be possible. This will allow increased stability of the therapeutic protein in human blood.

2.1.2.2 Genetic Engineering of the Microorganism: Techniques and Methodology

Genetic engineering is the production of new genes and alteration of genomes by substituting or adding new genetic material. The techniques of genetic engineering, alternatively known as recombinant DNA technology, made it possible to create unique industrial microorganisms able to synthesize valuable recombinant proteins. The basic series of steps for cloning a segment of foreign DNA to obtain a recombinant DNA molecule, as schematically explained in Figure 2.4, are:

1. DNA is purified from cells or tissues and the gene of interest is amplified by polymerase chain reaction (PCR).

2. DNA fragments, generated using enzymes called restriction enzymes, are ligated to other DNA molecules that serve as vectors, which has also been digested by the same restriction enzymes.

3. The recombinant DNA molecule consisting of the vector carrying an inserted foreign DNA fragment (perhaps a gene), is transferred to a host cell. Within the host cell, the recombinant molecule replicates, producing dozens of identical copies.

4. The clones carrying the correct form of the foreign DNA can be selected by plating the microorganisms in selective media, and the cloned DNA can be recovered from the host cells, purified and analyzed further.

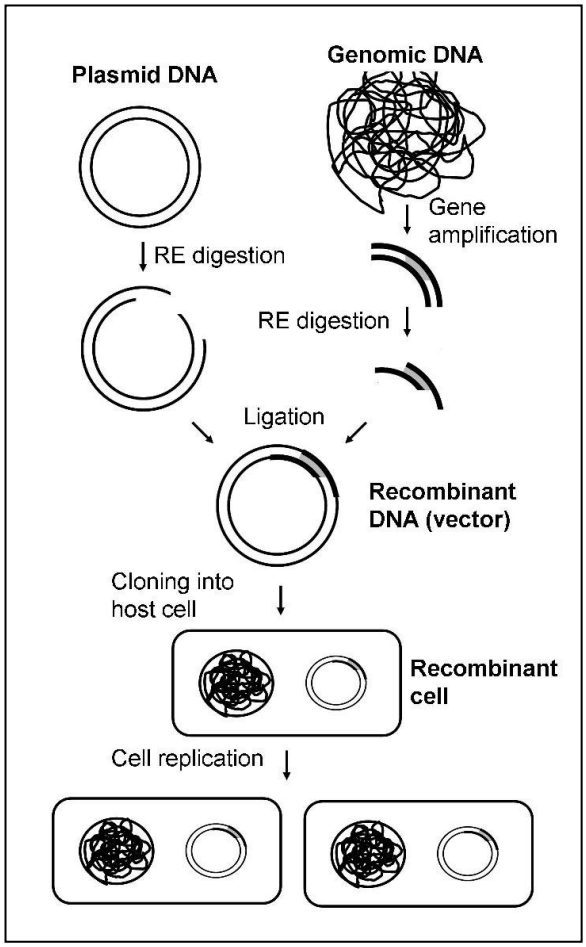


Figure 2.4 The basic series of steps for cloning a segment of foreign DNA to obtain a recombinant DNA molecule, the recombinant plasmid. The plasmid is then transferred to a host cell, which replicates producing more of the recombinant DNA.

In the first step of obtaining the gene of interest to be cloned (*eg. epo* in this study), the generally used technique is the polymerase chain reaction (PCR). Provided that, two short primer sequences (>20 nucleotides) on either side of the target gene is known, PCR is a useful technique for amplifying DNA sequences in vitro by separating the DNA into two strands and incubating it with oligonucleotide primers and DNA polymerase. It can amplify a specific sequence of DNA by as many as one billion times and is important in biotechnology, forensics, medicine, and genetic research. Three major steps are involved in a PCR device. These three steps are repeated for 30 or 40 cycles. The cycles are done on an automated cycler, which rapidly heats and cools the test tubes containing the reaction mixture. Each step - denaturation (the double-stranded DNA melts at greater than 90°C and opens into single-stranded DNA), annealing (joining primers and single stranded DNA), and extension (replication of DNA by polymerase at 72°C) - takes place at a different temperature (Figure 2.5). The annealing temperature is a key variable in determining the specificity of a PCR so temperatures and times used vary depending on the sequences to be amplified. The heat-stable enzyme commonly used is derived from a thermophilic Gram-negative eubacterium, *Thermus aquaticu*s (Glazer, 1995). One of these early isolates Taq DNA polymerase and its derivatives have a 5' to 3' polymerization depended exonuclease activity. For nucleotide incorporation, the enzyme works best at 75-80^{o}C, depending on the target sequence; its polymerase activity is reduced by a factor of 2 at 60^{o}C and by a factor of 10 at 37^{o}C (Sambrook and Russell, 2001).

Once the desired gene is isolated, it is inserted into a small piece of carrier DNA called a vector. Typically vectors are artificially constructed plasmids. A useful vector should: be able to replicate in the host cell; easily inserted into the host cell; contain a selectable marker (*eg.* drug resistance gene) to facilitate rapid, positive selection of cells which contain the vector; contains unique and convenient restriction sites for insertion of the DNA to be cloned. There are different kinds of yeast vectors: integrating plasmids, independently replicating plasmids and specialized plasmids. Most of these are shuttle plasmids, meaning that they permit DNA to be shuttled back and forth between yeasts and bacteria. Yeast integrative plasmids (*eg.* pPICZαA used in this study) are integrated into the yeast chromosome by homologous recombination and do not have a functional replication origins. It is advantageous in that the transformants are stable, even in the absence of the selective pressure (Walker, 1998).

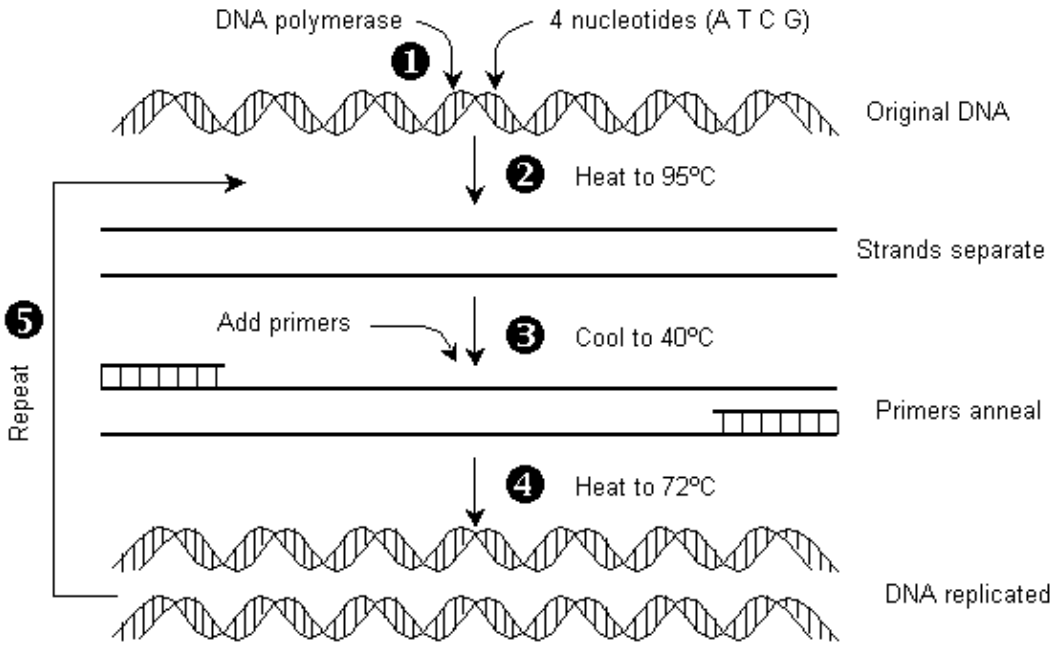


Figure 2.5 The Polymerase Chain Reaction (http://www.biologymad.com, Last Accesed: June 2008)

In order to produce recombinant plasmids, the gene of interest and the plasmid are digested with specific restriction enzymes prior to ligation. These enzymes naturally serve the bacteria as protection from foreign DNA. There are several types of restriction enzymes; the most useful for cloning, the Type II restriction enzymes recognize specific sequences, usually 4-8 bp in length, and cut DNA molecules within these sequences (Kirk and Othmer, 1994). The restriction enzymes used in this study and their recognition sequences are listed in Table 2.3.

Table 2.3 Recognition sites and cleavage points of restriction enzymes

Enzyme	Target site
*Eco R*I	5'-G^AATTCC-3'
Xba I	5'-T^CTAGA-3'
Sac I	5'-GAGCT^C-3'

The DNA fragment of interest cut by a restriction enzyme leads to single stranded tails, sticky ends, which have a tendency to anneal with the complementary strand present in the ligation reaction mixture. The addition of vector DNA cut open by the same restriction enzyme results in the annealing of the foreign DNA to the complementary ends of the cut vector. The phosphodiester bonds missing between the attached strands are covalently bond by DNA ligase. This enzyme catalyzes the condensation of 3'-hyroxyl group with a 5'-phosphate group to add the missing links. The ligation reaction is the rate limiting step in genetic engineering techniques since this reaction requires the cohesive ends of foreign DNA and open plasmid DNA to attach in correct orientation and anneal while preventing the religation of opened vector DNA.

After ligation, the mixture containing recombinant vector including the gene of interest is then transfered into the recipient or host cell. In most cases this is done by transformation (Schuler and Kargı, 2002). There are chemical, electrical and biolistic methods for transforming intact yeast cells (Walker, 1998). In this study, a chemical method was used where cells are permeabilized with a lithium salt (LiCl), DNA and PEG are added, and cells are briefly heat shocked.

Following transformation, it is important to note that construction of the desired vector-donor DNA usually results in a mixture including some opened or rejoined (without donor DNA) vector molecules, or insertion of DNA contaminants of donor DNA into the vector. Consequently, an efficient method to screen transformants for those with the desired vector-donor DNA combination is important (Schuler and Kargı, 2002). Antibiotic resistance is the most commonly used selection method because it allows for an extended host range and provides more flexibility in growth conditions (Smith, 1995). Thus, after transformation, cells possibly carrying the recombinant plasmid with the antibiotic resistance gene, are spread on plates containing a particular antibiotic to select the antibiotic resistant cells. The vector used in this study carries Zeocin resistance gene, which is an expensive but strong antibiotic. However, since the vector is a genomic integration vector, the antibiotic usage during fermentation is minimal and necessary only in the solid medium.

Not all of cells growing on the plates with specific antibiotic will carry the desired vector-donor DNA combination and candidates should be further analyzed. Candidates can be analyzed mainly by three methods: restriction mapping,

Southern blotting and DNA sequencing (Klug et al., 2006). All three methods have also been used in this study. A restriction map is a compilation of the number, order and distance between restriction enzyme-cutting sites along a cloned segment of DNA. Fragments generated by cutting DNA with restriction enzymes can be separated by agarose gel electrophoresis, which is a method that separates fragments by size, with the smallest pieces moving farthest. The fragments appear as a series of bands that can be visualized by staining the DNA with ethidium bromide and viewing under ultraviolet illumination. To make a Southern blot (after Edward Southern, who devised it) analysis, cloned DNA is cut into fragments with one or more restriction enzymes, and the fragments are separated by gel electrophoresis. The DNA in the gel is denatured into single-stranded fragments by treatment with an alkaline solution, and transferred to a sheet of DNA-binding membrane. The membrane-bound fragments are then hybridized with a labelled single-stranded DNA probe. The unbound probe is washed away, and the hybridized fragments are visualized on a piece of film. The ultimate way to characterize a clone is DNA sequencing, because it can detect even single base pair mutations, but is the most complex of the three methods. Fortunately, technology today employs automated DNA sequencers that can sequence whole genomes in as short as one day.

2.1.3 Medium Design and Bioreactor Operation Parameters

The productivity of a recombinant system depends on several genetic and physiological factors such as the codon usage of the expressed gene, the gene copy number, efficient transcription by using strong promoters, translation signals, translocation determined by the secretion signal peptide, processing and folding in the endoplasmic reticulum and Golgi and, finally secretion out of the cell, as well as protein turnovers by proteolysis, but also on the optimization of fermentation strategy (Cos et al, 2006). In terms of fermentation strategy, the important criteria that must be taken into account are:

1. Medium design
2. Bioreactor operation parameters
    - pH
    - Temperature
    - Oxygen transfer rate

2.1.3.1 Medium Design

Conceptually, the cells function as semi-batch microbioreactors with volume V, wherein the biochemical reactions take place (Çalık et al., 2003). Between the metabolic system and the bioreactor medium, which is defined as the environment, a number of compounds of the intracellular biochemical reaction network (e.g substrate(s), oxygen, H^+, H_2O, CO_2, amino acids, organic acids of the glycolysis or the gluconeogenesis pathways and the tricarboxylic acid (TCA) cycle) are exchanged or transferred with facilitated and active transport mechanisms. Furthermore, most of the products formed by organisms are produced as a result of their response to environmental conditions, such as nutrients, growth hormones, and ions. The qualitative and quantitative nutritional requirements of cells need to be determined to optimize growth and product formation (Shuler and Kargı, 2002). Typical components of a fermentation medium are explained below:

1. *Water* is the major component of all fermentation media.
2. *Macronutrients* are needed in concentrations larger than 10^{-4}M. Carbon, nitrogen, oxygen, hydrogen, sulphur, phosphorus, Mg^{2+}, and K^+ are major macronutrients.
3. *Micronutrients* are needed in concentrations of less than 10^{-4}M. Trace elements such as Mo^{2+}, Zn^{2+}, Cu^{2+}, Mn^{2+}, Fe^{2+}, Ca^{2+}, Na^{2+}, vitamins, growth hormones, and metabolic precursors are micronutrients.
4. *Buffers* may be necessary to control the pH of the fermentation medium.
5. *Vitamins and growth factors* are required in some fermentation processes.
6. *Inducers* must be present in the medium if the product of interest is synthesized in response to the presence of an inducer in the environment.
7. *Antifoams* are surface active agents, reducing the surface tension in the foams. They usually have no metabolic effect but may substantially reduce the oxygen transfer rate.

There are two major types of growth media; defined medium, containing specific amounts of pure chemical compounds with known chemical compositions, and complex medium, containing natural compounds whose chemical composition is not exactly known. Complex medium often results in higher cell and protein yields due to its rich ingredients. On the other hand, defined medium allows better control over the fermentation and leads to easier and cheaper recovery and purification of the product (Shuler and Kargı, 2002).

Certainly, there are not fixed conditions that ensures the optimal production, but there are several guidelines that allow significant improvement of the productivity.

The most common medium for high cell density fermentation of methylotrophic yeast *P. pastoris* is the basalt salt medium (BSM) along with its companion trace salts medium (PTM1). This is considered a standard medium, though it may not be the optimum and may have some important problems, such as unbalanced composition, precipitates and high ionic strength (Cereghino et al., 2002; Cos et al., 2006). Thus, modifications have been tried in several studies (Brady et al., 2001; Thorpe et al., 1999; Jungo et al., 2006) and the effect of each medium component have been investigated in detail by Plantz et al. (2007). The nitrogen source in this standard medium is mainly added as ammonium hydroxide solution when controlling pH (Cos et al., 2006). This method has the advantage of avoiding nitrogen accumulation, which can provoke the inhibition of growth and enlarge the lag phase (Yang et al., 2004). The disadvantage arises in the case of inadequate feeding of nitrogen source, causing nitrogen starvation and consequently increased protease secretion.

Carbon source plays an important role over recombinant protein production and cell growth. Most commonly used carbon sources for *Pichia pastoris* are methanol, glycerol, sorbitol, glucose, mannitol, trehalose, etc (Brierley et al., 1990; Sreekrishna et al., 1997; Thorpe et al., 1999; Inan and Meagher, 2001-b). Methanol is used not only as an inducer for the expression of recombinant protein, but also as the carbon and energy source. However, above certain concentrations, growth is substrate-inhibited by methanol (Zhang et al., 2000b); therefore a fed-batch protocol is generally used.

The standard protocol for the expression of recombinant proteins in *P. pastoris*, under the control of the *AOX*1 promoter, is usually performed in three stages (Stratton et al., 1998). First, cells are grown in batch defined medium containing glycerol in order to achieve high cell densities rapidly, while repressing foreign gene expression since the *AOX*1 promoter is repressed by unlimited growth on glycerol (Tschopp et al., 1987). Secondly, glycerol is fed at limiting concentrations in order to increase biomass concentration further. This transition phase allows derepressing the enzymes necessary for the dissimilation of methanol gradually and reduces the time necessary for the cells to adapt to growth on methanol (Chiruvolu et al., 1997). Finally, recombinant protein production is induced by the addition of methanol in fed-batch mode.

Since *Pichia pastoris* is sensitive to methanol concentration, it is an important process parameter, and methanol needs to be supplied continuously to the growing culture, while keeping its concentration below toxic limits, 4 g L^{-1}. Thus, several other fed-batch strategies for methanol addition have been established (Zhang et al., 2000b). These strategies can be metabolism related, based on parameters such as methanol consumption (Curvers et al., 2001-a; Guarna et al., 1997; Hellwig et al., 2001; Katakuga et al., 1998; Kobayashi et al., 2000-a), oxygen consumption (Byrne et al., 2000; Chung, 2000; Minning et al., 2001), CO_2 concentration or pH control. Growth strategies can also be based on predetermined methanol feeding schemes at constant, linear, or exponential rates (Chauhan et al., 1999; Freyre et al., 2000; Inan et al., 1999; Murasugi et al., 2000, Zhang et al., 2000a), where the feeding rate determines the growth rate. A comparative study was performed by Trinh et al. (2003), where various methanol feeding strategies were investigated for increased cell and recombinant protein productivities. Feeding methanol at a predetermined exponential profile and controlling the growth rate at 0.02 h^{-1} was found to be more efficient than the other two strategies, which were using an online methanol sensor or spike of dissolved oxygen signal for controlling methanol feed rates.

The use of multi-carbon substrates in addition to methanol is the other strategy to increase cell density and process productivity, as well as to reduce the induction time. This strategy has been mostly employed for fermentations using Mut^s strains because of their genetically reduced capacity to assimilate methanol (Ramon et al., 2007). Nevertheless, increased attention on mixed substrate growth can be seen for Mut^+ strains as well, due to its various advantages.

The use of mixed feeds of methanol and glucose with the methylotrophic yeasts (*Hansenula polymorpha* and *Candida boidinii)* was first investigated in the 1980's by Egli and coworkers (Egli et al., 1980; Egli et al., 1982b; Egli et al., 1982a; Egli et al., 1983; Egli et al., 1986). It was shown that during mixed substrate growth, methanol was utilized at dilution rates significantly higher than the maximum specific growth rate achieved on methanol as sole carbon source. Hence, higher productivities can be achieved with mixed substrates than with methanol as sole carbon source (Egli et al., 1982a). However, derepression of methanol dissimilating enzymes was observed during glucose-limited continuous cultures (Egli et al., 1980). Although the extent to which derepression occurred was different for each enzyme, higher specific enzyme activities in the cells were observed at low dilution rates for all enzymes. It was pointed out that since the residual substrate concentration decreases with decreasing dilution rates in chemostat cultures, the extent of the observed derepression most likely relates to the residual concentration of glucose.

Brierley et al. (1990) was the first group attempting fed-batch strategy using mixed substrates of glycerol and methanol for *Pichia pastoris*. This strategy was followed by several other groups, to increase the volumetric protein productivity as a result of higher cell densities and feeding rates possible with growth on glycerol (Cregg et al., 1993; Loewen et al., 1997; McGrew et al., 1997; Katakura et al., 1998; Zhang et al., 2003-a). However, some authors claim that the optimal level of protein expression is not achievable with mixtures of glycerol and methanol, due to a partial repression of the *AOX*1 promoter by glycerol, which may result in lower specific productivities of recombinant protein (Sreekrishna et al., 1997; Hellwig et al., 2001; Xie et al., 2005).

On the other hand, sorbitol is a non-repressing carbon source for *AOX1* promoter, thus sorbitol accumulation during the induction phase does not affect the expression level of recombinant protein. These advantages, although widely utilized for Mut^S strains, have recently been appreciated for Mut^+ strains and investigated in a limited number of studies. The idea was first recommended in the study of Sreekrisha et al. (1997), however first experimental data, showing the advantage of sorbitol supplement for the induction phase of Mut^+ strains, was given by Inan et al. (2001-b), in one batch shake-bioreactor experiment. Thereafter, Ramon et al. (2007) commented on the consumption mechanism of sorbitol and methanol, from a batch bioreactor experiment, that these substrates

were consumed sequentially. Finally, Jungo et al. (2007) optimized the sorbitol content in the feed in a single continuous bioreactor experiment, and thereafter at this feeding ratio, performed two fed-batch bioreactor experiments, at μ=0.03 h^{-1} and μ=0.05 h^{-1}. The main conclusions driven in this study were on simultaneous consumption of methanol and sorbitol (conflicting with the study of Ramon et al., 2007); and that at μ=0.05 h^{-1}, accumulation of sorbitol did not affect the specific productivity.

Thus, the present studies in the literature have not fully investigated the fed-batch operational strategies in the presence of sorbitol, the effect of sorbitol on cell metabolism of the Mut^+ *P. pastoris* strain, nor have formed any mathematical models to explain the production system. Moreover, since sorbitol is a non-repressing carbon source, there is no need to complicate the bioprocess by feeding the sorbitol fed-batch together with methanol. Thus, sorbitol can easily be implemented to the existing processes, simply by adding the sorbitol in batch mode at the induction phase. It is not meaningful to modify the first batch stage since the maximum specific growth rate on sorbitol is μ_{max}= 0.032 h^{-1}, whereas on glycerol it is μ_{max}= 0.18 h^{-1}.

Additional advantages of mixed substrate growth on sorbitol and methanol are significant reductions in heat production and oxygen consumption rates. Indeed, since the enthalpy of combustion of sorbitol is about 8% lower than that of glycerol and about 30% lower than that of methanol, for a given growth rate, less heat will be released in cultures with mixed carbon sources of sorbitol and methanol than with mixed carbon sources of glycerol and methanol or with methanol alone. Moreover, since the degree of reduction of sorbitol is lower than those of glycerol and methanol, less oxygen will be consumed during mixed substrate growth on sorbitol and methanol than on mixed feeds of glycerol and methanol or on methanol as sole carbon source (Jungo et al., 2007).

2.1.3.2 pH

Hydrogen ion concentrations (pH) affect the activity of enzymes, transport mechanisms and other extracellular and intracellular events, thus the microbial growth rate. Microbial cells have a remarkable ability to maintain the intracellular pH at a constant level, even with large variations in the pH of the extracellular medium, but only at the expense of an increase in the maintenance energy

demands, since Gibbs free energy has to be used for maintaining the proton gradient across the cell membrane. The influence of the bioreactor operation parameter pH on the bioreaction network is indeed important, and needs clarification in order to develop an efficient bioreactor operation strategy.

P. pastoris is capable of growing across a relatively broad pH range (3.0–7.0). This range does not affect the growth significantly, which allows considerable freedom in adjusting to a pH favourable for recombinant protein production and stability (Macauley-Patrick et al., 2005). Different pH values were found to be optimal depending on the nature of the recombinant protein and its stability: pH 6.0 was optimal in production of recombinant mouse epidermal factor and human serum albumin (Clare et al., 1991; Kobayashi et al., 2000-b) and pH 3.0 was optimal in production of insulin-like growth factor-I and cytokine growth-blocking peptide (50 mg L^{-1}) (Brierley et al., 1994; Koganesawa et al., 2002). Most commonly though, the pH value has been fixed around 5.5 to reduce protease effects in the medium, and to improve the stability of the foreign protein (Cos et al., 2006).

2.1.3.3 Temperature

Temperature is one of the most important bioprocess parameters which is normally desired to be kept constant at its optimum value throughout the bioprocess (Nielsen et al., 2003). Microorganisms do not have the ability to regulate their internal temperature. Thus, cell temperature is always equal to environmental temperature and all biochemical reaction rates depend directly on the external temperature. It may affect both the growth rate and the product formation; however, the optimum temperature for growth and product formation may be different. The influence of temperature on the maximum specific growth rate of a microorganism is similar to that observed for enzyme activity: an increase with increasing temperature up to a certain point where protein denaturation starts, and a rapid decrease beyond this point. For temperatures below the onset of protein denaturation the maximum specific growth rate increases much the same way for a normal chemical rate constant, explained by Arrhenius equation (Nielsen et al., 2003). On the other hand, when temperature is increased above the optimum temperature, the maintenance requirements of cells will increase. Moreover, metabolic regulations, nutritional requirements,

biomass composition and the yield coefficients will also be affected by temperature.

Most processes with *P. pastoris* are run at an optimum temperature for growth of 30°C (Wegner, 1983). Though, few studies have been carried out on the effect of temperature on recombinant protein production in *P. pastoris*. It has been stated that temperatures above 32°C can be detrimental to protein expression and may lead to cell death (Invitrogen, 2002), and temperatures above 30°C were not appropriate for the production of a recombinant peptide (Inan et al., 1999), since elevated temperatures result in cell death, which will lead to cell lysis and higher protease activity in fermentation media.

On the other hand, lower cultivation temperature generally influences the yield of recombinant protein produced by *Pichia pastoris* because protease release to the medium is reduced due to lower cell death, decreasing the degradation of proteins (Macauley-Patrick et al., 2005). Expression of recombinant proteins at lower temperature also helps to reduce protein misfolding and to produce more properly folded proteins to be secreted into the culture medium (Georgiou and Valax, 1996). The increased yields of recombinant proteins at lower temperatures have been reported in many studies (Chen et al., 2000; Whittaker and Whittaker, 2000; Li et al., 2001; Hong et al., 2002; Sarramegna et al., 2002; Jahic et al., 2003a; Jahic et al., 2003b; Li et al., 2003; Shi et al., 2003). However, it has to be pointed out that there many other cases where lowering the temperature below 30°C does not significantly influence the production of recombinant proteins expressed by *P. pastoris* strains (Inan et al., 1999; Curvers et al., 2001-b; Hong et al., 2002; Kupesulik and Sevella, 2005).

2.1.3.4 Oxygen Transfer Rate

Oxygen shows diverse effect on product formation in aerobic fermentation processes by influencing metabolic pathways and changing metabolic fluxes (Çalık et al., 1999). Molecular oxygen is used as a terminal electron acceptor in the aerobic metabolism of carbon compounds. Gaseous oxygen is introduced into the growth media by sparging air or by surface aeration, thus oxygen transfer rate can be adjusted by either changing the air inlet rate or agitation rate.

One of the reasons for the popularity of the *P. pastoris* expression system is that physiologically it prefers a respiratory rather than a fermentative mode of growth (Cereghino et al., 2002), thus, it does not produce inhibiting products such as ethanol and acetic acid. It is an obligately aerobic organism when it grows on methanol and it requires high oxygen transfer rates throughout the methanol metabolism. The necessity for aeration with pure oxygen for the achievement of high cell densities with *P. pastoris* cultures grown on methanol is one of the drawbacks of this expression system. On the other hand, there are studies where oxygen limited fed batch processes were used (Trentmann et al., 2004; Trinh et al., 2003). Charoenrat et al, (2005) compared oxygen limited fed batch (OLFB) process with methanol limited fed batch (MLFB) process for the production of Thai Rosewood β-glucosidase and they found 35% higher oxygen uptake rate, higher productivity and specific activity in OLFB when compared with MLFB. Nevertheless, majority of the studies carried out with *P. pastoris* use supplementation with pure oxygen to keep the dissolved oxygen level above 20 – 30% during the whole induction phase on methanol (Jahic et al., 2006).

2.1.4 Recombinant Protein Purification

Once a recombinant protein is produced, whether for structural analysis or for meeting the regulatory standards in marketing the product, the product needs to be purified. Most of the recent developments in the area of protein bioseparations are centered on therapeutic proteins, which require high purity (99.9%) and well-characterized structure. The key considerations in purification are, keeping it simple by minimizing the number steps, employing fast, easy and cheap processes that will not denature the protein. Due to the high costs required for product recovery steps, many fermentation strategies can also be applied, such as using a defined chemical growth medium rather than a complex medium and choosing a strain which produces an extracellular rather than an intracellular product (Bailey and Ollis, 1986). The product was designed to be produced extracellularly in this study.

The typical operations sequence through which a bioreactor broth must pass for a highly purified product is given in Table 2.4. The underlined techniques are the ones used in this study.

Table 2.4 Typical stages and techniques used in protein purification (Bailey and Ollis, 1986; Roe, 2001). Techniques used in this study are underlined.

Stage	Aim	Typical Techniques
1. Initial fractionation	Clarification of the protein solution, removal of cell debris	• Centrifugation • Filtration • Sedimentation
2. Primary isolation	Concentration and removal of contaminants	• Precipitation • Solvent extraction • Ultrafiltration
3. Purification	Further product concentration and purification from other proteins	• Affinity separation • Ion exchange Chromatography • Size exclusion Chromatography
4. Polishing	Final formulation to prepare for storage/shipping	• Drying • Crystallization

Affinity separation is a powerful method for separating the desired protein from the others by using the specific binding property of proteins. Moreover, affinity purification tags can be fused to any recombinant protein of interest, allowing fast and easy purification following a procedure that is based on the affinity properties of the tag rather than the protein (Nilsson et al. 1997). Many different affinity tags have been developed to simplify protein purification. These affinity-tag systems share the following features: (a) one-step adsorption purification; (b) a minimal effect on tertiary structure and biological activity; (c) easy and specific removal to produce the native protein; (d) simple and accurate assay of the recombinant protein during purification; (e) applicability to a number of different proteins (Terpe, 2003).

The most commonly used tag to purify and detect expressed recombinant proteins is the polyhistidine tag (Yip et al. 1989). Protein purification using polyhistidine tags relies on the affinity of histidine residues for immobilized metal such as nickel or cobalt, which allows selective protein purification (Yip et al.

1989; Hutchens and Yip, 1990). The principle is based on the reversible interaction between the histidine side chains and immobilized metal ion. Under conditions of physiological pH, histidine binds by sharing nitrogen electron density with the electron-deficient orbital of the metal ion. After purification of the recombinant protein, it is separated from the resin by adding imidazole to the elution buffer, because imidazole is identical to the histidine side chain.

Once the protein has been isolated, it is usually necessary to remove the tag before subsequent use of the expressed protein of interest (Araujo et al., 2000). The most popular method to remove the tag involves the use of a specific protease such as thrombin or factor Xa. This involves the insertion of a unique amino acid sequence that is specific for cleavage by either factor Xa protease or thrombin between the protein of interest and the fusion partner (Jenny et al., 2003).

2.1.5 Structural Analysis of the Recombinant Protein

After successful cloning, production and purification of the recombinant protein, the product also has to prove the right structure. There are many different methods for the primary structural analysis of the recombinant proteins produced. In this study, to determine the presence of the recombinant protein and its molecular weight, three typically used techniques were employed, which are SDS-PAGE, western blotting and MALDI-ToF MS.

2.1.5.1 SDS-Polyacrylamide Gel Electrophoresis

This standard method of protein analysis for rough determination of the size of the polypeptides, is carried out in polyacrylamide gels under conditions that ensure dissociation of proteins into their individual polypeptide subunits and that minimize aggregation. Most commonly, the strong anionic detergent SDS (sodium dodecyl sulfate) is used in combination with a reducing agent (usually 2-mercaptoethanol) and heated to dissociate the proteins before they are loaded onto the gel. The denatured polypeptides bind SDS and become negatively charged. Because the amount of SDS bound is almost always proportional to the molecular weight of polypeptide and is independent of its sequence, SDS-polypeptide complexes migrate through polyacrylamide gels in accordance with the size of polypeptide. Unlabeled proteins separated by polyacrylamide gel electrophoresis typically are detected by staining, either with Coomassie Brilliant

Blue or with silver salts. In relatively rapid and straightforward reactions, Coomassie Brilliant Blue binds nonspecifically to proteins but not gel, thereby allowing visualization of proteins as discreet blue bands within the translucent matrix of the gel. Silver staining is significantly more sensitive, although somewhat more difficult to perform. The use of silver staining allows detection of proteins resolved by gel electrophoresis at a concentrations nearly 100-fold lower than those detected by Coomassie Brilliant Blue staining (Sambrook and Russel, 2001).

2.1.5.2 Western Blotting

Western blot analysis is used to detect the presence of a specific protein in a mixture of proteins while giving information about the size of the protein as well. In this technique the antibody-reactive proteins in a mixture are analyzed by first resolving the proteins in that mixture by SDS-PAGE. After electrophoresis, the gel is placed in contact with a nitrocellulose membrane, and the proteins are transferred to the membrane by electric current. The proteins are bound irreversibly to the membrane, so the antigen-antibody reactions can be visualized after treatment of the membrane with antibody and the substrate of the enzyme attached to the antibody. The enzyme-substrate complex generally gives a colored or chemiluminescent signal for detection.

2.1.5.3 MALDI-ToF Mass Spectrometry

A more reliable and sensitive mass analyzer used in protein analysis is the matrix-assisted laser desorption/ionization time-of-flight (MALDI-ToF) mass spectrometry. The sample which is dispersed in a large excess of matrix material strongly absorbs the incident light. Short pulses of laser light focused on to the sample spot causes the sample and matrix to volatilize. The ions formed are accelerated by a high voltage supply and then allowed to drift down a flight tube where they are separated according to mass. Arrival at the end of the flight tube is detected and recorded by a high speed recording device (Figure 2.6).

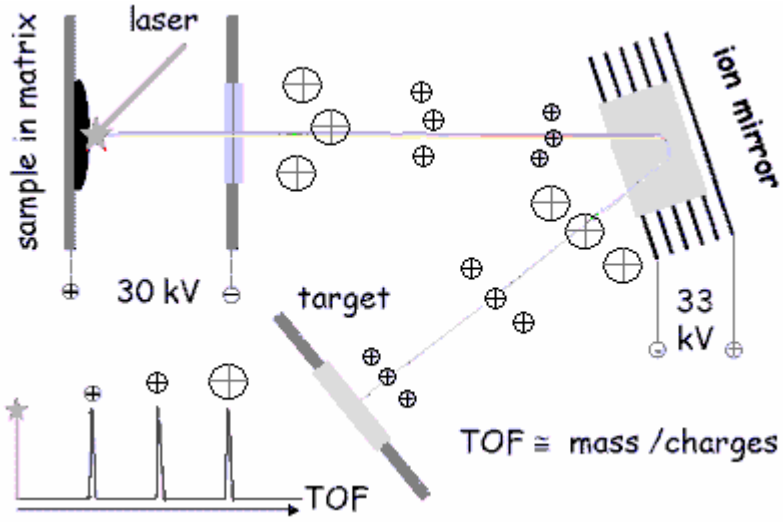


Figure 2.6 Schematic representation of MALDI-ToF MS working principle in reflectron mode.

An accelerating potential (V) will give an ion of charge z an energy of zV. This can be equated to the kinetic energy of motion, and the mass (m) and the velocity (v) of the ion;

$$zV = 1/2mv^2 \qquad (2.1)$$

Since velocity is length (L) divided by time (t) then,

$$m/z = [2Vt^2]/L^2 \qquad (2.2)$$

V and L cannot be measured with sufficient accuracy but the equation can be rewritten as,

$$m/z = B(t-A)^2 \qquad (2.3)$$

where, A and B are calibration constants that can be determined by calibrating to a known m/z. Thus, x-axis of the spectrum will be m/z. Molecules in the spectrum will usually be seen as isotopic clusters, separated by 1 unit on the *m/z* axis, which means z = 1. Therefore, mass (m) of the molecule can be easily calculated.

2.2 Computation of Bioprocess Characteristics

2.2.1 Yield Coefficients and Specific Rates

A simple but important step during a bioprocess is the computation of the yield coefficients and the specific rates. Stoichiometrically related parameters, namely yield coefficients are commonly used to better evaluate the bioprocess. As a general definition,

$$Y_{P/S} = -\frac{\Delta P}{\Delta S} \tag{2.4}$$

where, $Y_{P/S}$ is the overall yield coefficient, P and S are product and substrate, respectively, involved in metabolism. ΔP is the mass or moles of P produced, and ΔS is the mass or moles of S consumed. This definition gives an overall yield representing some sort of average value for the entire culture period. However, in batch and fed-batch processes, the yield coefficients may show variations throughout the process for a given microorganism in a given medium, due to the growth rate and metabolic functions of the microorganism. Therefore, it is sometimes necessary to evaluate the instantaneous yield at a particular point in time. Instantaneous yield can be calculated as follows:

$$Y'_{P/S} = -\frac{dP}{dS} = -\frac{dP/dt}{dS/dt} \tag{2.5}$$

When yields for fermentation are reported, the time or time period to which they refer should be stated (Doran, 1995). A list of frequently used yield coefficients is given in Table 2.5.

The formation rates of biomass, metabolic products and consumption rate of substrate can be determined from measurements of the corresponding concentrations. Often it is convenient to normalize the rates with respect to the amount of biomass present, since the rates can then be easily compared between fermentation experiments, even if the amount of biomass changes. Such normalized rates are referred to as specific rates.

Table 2.5 Definition of yield coefficients.

Symbol	Definition	Unit
$Y_{X/S}$	Mass of cells produced per unit mass of substrate consumed	kg cell kg^{-1} substrate
$Y_{X/O}$	Mass of cells produced per unit mass of oxygen consumed	kg cell kg^{-1} oxygen
$Y_{S/O}$	Mass of substrate produced per unit mass of oxygen consumed	kg substrate kg^{-1} oxygen
$Y_{P/X}$	Mass of product formed per unit mass of substrate consumed	kg product kg^{-1} cell
$Y_{P/S}$	Mass of product formed per unit mass of substrate consumed	kg product kg^{-1} substrate

The specific growth rate (μ) is a very important process variable. For fed-batch processes from general mass balance for biomass,

$$r_X V = \frac{d(C_X V)}{dt} \tag{2.6}$$

where, the main assumption is that the cells are batch-wise and are not lost through sampling. The biomass formation rate, r_X, is defined as the product of specific cell growth rate (μ), and cell concentration (C_X), i.e.,

$$r_X = \mu C_X \tag{2.7}$$

Combining equations (2.6) and (2.7),

$$\frac{d(C_X V)}{dt} = \mu C_X V \tag{2.8}$$

Volume change is due to methanol feed with a predetermined rate, thus for simplicity, system is assumed to be constant density, and from overall mass balance,

$$\frac{dV}{dt} = Q \tag{2.9}$$

Thus, from cell mass concentration (C_X), volume (V) data throughout the bioprocess, μ can be calculated easily by inserting equation (2.8) into (2.9),

$$\frac{dC_X}{dt} = \left(\mu - \frac{Q}{V}\right) C_X \tag{2.10}$$

and then rearranging equation (2.10),

$$\mu = \frac{dC_X}{dt}\frac{1}{C_X} + \frac{Q}{V} \tag{2.11}$$

Mass balance for the substrate given batch-wise to a fed-batch system (sorbitol in this study) and thus C_{S0}=0, would be

$$r_s V = \frac{d(C_s V)}{dt} \tag{2.12}$$

Similar to equation (2.7),

$$r_S = q_S C_X \tag{2.13}$$

Inserting equation (2.13) into (2.12),

$$\frac{dC_S}{dt} = -\frac{Q}{V}C_S + q_S C_X \tag{2.14}$$

and rearranging equation (2.14),

$$q_S = \left(\frac{dC_S}{dt} + \frac{Q}{V}C_S\right)\frac{1}{C_X} \tag{2.15}$$

Mass balance for the fed-batch substrate, (methanol in this study), added with an exponential volumetric flow rate of $Q_,$ in fed-batch process with a reaction volume V, can be written as

$$QC_{Mo} - 0 + r_M V = \frac{d(C_M V)}{dt} \tag{2.16}$$

The substrate consumption rate (r_M) can be defined as the product of specific substrate consumption rate (q_M) and cell concentration (C_X), i.e.,

$$- r_M = q_M C_X \tag{2.17}$$

Inserting equation (2.17) into (2.16)

$$QC_{M0} - q_M C_X V = V\frac{dC_M}{dt} + C_M\frac{dV}{dt} \tag{2.18}$$

A fed-batch system is generally assumed to operate at quasi-steady state when nutrient consumption rate is nearly equal to nutrient feed rate. Thus, at quasi-steady state, no significant level of the substrate (methanol) can accumulate (Shuler and Kargı, 2002). Therefore, although the terms on the right hand side are physically and generally negligible, mathematically the $V(dC_M/dt)$ term should be considered carefully for each specific case. The last term can be neglected though,

$$C_M \frac{dV}{dt} \sim 0 \tag{2.19}$$

Combining equations (2.16) - (2.19), dividing by V and rearranging, the general mass balance simplifies to,

$$\frac{dC_M}{dt} = \frac{Q}{V}(C_{M0}) - q_M C_X \tag{2.20}$$

Thus, specific methanol consumption rate, q_M can be calculated after rearranging equation (2.20), where again the last term on right hand side is generally negligible in quasi-steady states,

$$q_M = \frac{Q}{V}\frac{C_{M0}}{C_X} - \frac{1}{C_X}\frac{dC_M}{dt} \tag{2.21}$$

From a similar mass balance given in equation (2.14), the recombinant protein balance results in,

$$\frac{dC_{rp}}{dt} = q_{rp} C_X - \frac{Q}{V} C_{rp} \tag{2.22}$$

Thus, after rearrangement, the specific recombinant protein production rate (q_{rp}) can be calculated using experimental data from,

$$q_{rp} = \left(\frac{dC_{rp}}{dt} + \frac{Q}{V} C_{rp} \right) \frac{1}{C_X} \tag{2.23}$$

In literature, the term "QC/V" is referred to as the extracellular dilution term. If this term is omitted, the fed-batch process would be oversimplified to be a batch process.

In summary, the specific rates μ, q_S, q_M and q_{rp} can be calculated from experimental data using equations (2.11), (2.14), (2.21), (2.23), respectively.

2.2.2 Oxygen Transfer Characteristics

Pichia pastoris is known to require much more oxygen then a regular yeast, due to its methanol metabolism. Thus, oxygen transfer characteristics are highly important for this yeast.

The transfer of oxygen from the fermentation medium to microorganism takes place in several steps (Bailey and Ollis, 1986). When cells are dispersed in the liquid, and the bulk fermentation broth is well mixed, the major resistance to oxygen transfer is the liquid film surrounding the gas bubbles; therefore the rate of oxygen transfer from gas to liquid is of prime importance. An expression for oxygen transfer rate (OTR) from gas to liquid is given by,

$$OTR = K_L a(C_O^* - C_O) \tag{2.24}$$

where, C_O^* is saturated dissolved oxygen concentration and C_O is the actual dissolved oxygen concentration in the broth and a is the gas-liquid interfacial area. Since solubility of oxygen in aqueous solutions is very low, the liquid phase mass transfer resistance dominates, and the overall liquid phase mass transfer coefficient, $K_L a$, is approximately equal to liquid phase mass transfer coefficient, $k_L a$ (Shuler and Kargı, 2002). Moreover, the maximum possible mass transfer rate is defined as,

$$OTR_{\max} = K_L a C_O^* \tag{2.25}$$

On the other hand, oxygen uptake rate (OUR), $-r_O$, per unit volume of broth is given by

$$OUR = -r_O = q_O C_X \tag{2.26}$$

where, q_O is the specific rate of oxygen consumption and C_X is the cell concentration (Shuler and Kargı, 2002).

Numerous methods have been developed for the experimental determination of $K_L a$ values. The dynamic method (Bandyopadhyay and Humprey, 1967) is a widely used method and can be applied during the fermentation process. The method is based on a material balance for oxygen given by,

$$\frac{dC_O}{dt} = OTR - OUR = K_L a(C_O^* - C_O) + r_O \tag{2.27}$$

To apply the method, as shown in Figure 2.7, at some time t_0, the broth is first de-oxygenated by stopping the air flow and lowering the agitation rate to a minimum level to prevent surface aeration. During this period, dissolved oxygen concentration (C_O) drops due consumption by the microorganism, and since there is no oxygen transfer, equation (2.27) becomes:

$$\frac{dC_O}{dt} = r_O \tag{2.28}$$

Using equation (2.28) in region-II of Figure (2.7), oxygen uptake rate, $-r_O$, can be determined.

Air inlet is then turned back on, and the increase in C_O is monitored as a function of time. In this period, region-III, equation (2.27) is valid, and r_O calculated from eqauation (2.28) in region-II is assumed to have the same value in region-III for the same analysis set, since one analysis takes about 10 min. Thus, rearranging equation (2.27),

$$C_O = -\frac{1}{K_L a}\left(\frac{dC_O}{dt} - r_O\right) + C_O^* \tag{2.29}$$

From the slope of a plot of C_O versus ($dC_O/dt - r_O$), K_La can be determined, as demonstrated in Figure 2.8.

The Dynamic Method can also be applied to conditions under which there is no reaction (no microorganism in the medium), i.e., $r_O=0$ (Nielsen et al., 2003). In this case, the broth is de-oxygenated by sparging nitrogen into the vessel. Air inlet is turned back on and again the increase in C_O is monitored as a function of time. Modifying equation (2.29)

$$C_O = -\frac{1}{K_L a}\frac{dC_O}{dt} + C_O^* \tag{2.30}$$

From the slope of a plot of C_O versus $d(C_OV)/dt$, the physical mass transfer coefficient, K_La_0, can be determined.

In order to compare the relative rates of maximum oxygen transfer and biochemical reactions and find the rate limiting step of the bioprocess, the maximum possible oxygen utilization rate (OD=oxygen demand) which is defined as (Çalık et al., 2000),

$$OD = \frac{\mu_{\max} C_X}{Y_{X/O}} \tag{2.31}$$

In order to express the oxygen limitation in an aerobic process, the effectiveness factor, η (oxygen uptake rate per maximum possible oxygen utilization rate) and Damköhler number, Da (maximum possible oxygen utilization rate per maximum mass transfer rate) are defined according to the equations given below (Çalık et al., 2000):

$$\eta = \frac{OUR}{OD} \tag{2.32}$$

$$Da = \frac{OD}{OTR_{\max}} \tag{2.33}$$

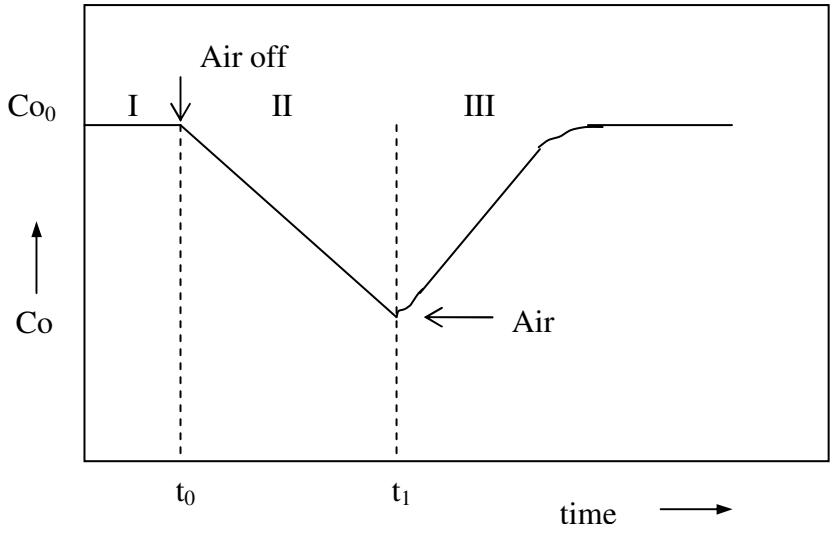


Figure 2.7 Variation of dissolved oxygen concentration with time in dynamic measurement of $K_L a$.

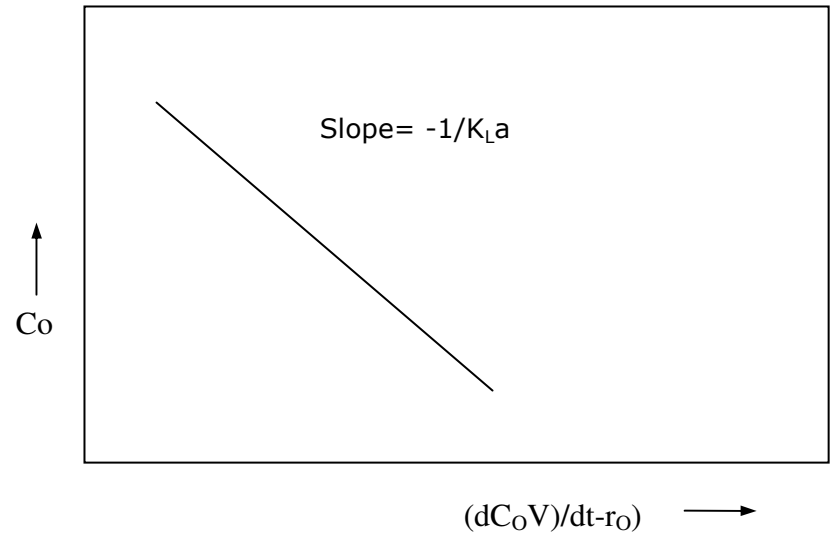


Figure 2.8 Evaluting $K_L a$ using the Dynamic Method.

2.3 Mathematical Modelling of Metabolism

Mathematical models can serve as a guide in choosing among different possible regulatory structures for a specific cellular process, and thus can be used to determine the maximum production capacity of the microorganism. The currently used metabolic modelling approaches can be subdivided according to their structure as stoichiometric models and kinetic models (Gombert and Nielsen, 2000).

2.3.1 Stoichiometric Models

Metabolic flux analysis (MFA) is a widely used tool for stoichiometric modeling of cellular metabolites and is based on calculation of intracellular reaction network rates through various reaction pathways. MFA is required for the quantitative analysis of the cell metabolism and is helpful to:

- calculate the intracellular reaction rates
- calculate the theoretical metabolic capacities of the microorganism
- modify the medium composition
- improve the bioreactor operation conditions
- identify the existence of different pathways
- calculate non-measured fluxes
- examine the influence of alternative pathways
- elucidate the physiological state of the cells,
- select the host microorganism for recombinant biomolecule production (Çalık and Özdamar, 2002-a).

Initially, an idea of the main biochemical processes that occur inside the organism being studied and the chemical reactants involved are needed. A biochemical reaction network is formed, starting from the site that the substrate(s) enters into the carbon metabolism, and all the other relevant bioreactions of the central carbon metabolism should be considered; i.e., the glycolysis and the gluconeogenesis pathways, the pentose phosphate pathway (PPP), the tricarboxylic acid (TCA) cycle, and the anaplerotic reactions. The

overall pathway can be simplified based on a comparison of the extracellular growth compounds and the intracellular chemical compounds, by lumping some reactions into single ones without loosing representation accuracy (Çalık and Özdamar, 2002-a). Thus, a key issue in the design of novel fermentation processes, and in the optimization of existing fermentation processes is a quantitative description of the cellular reactions that play an important role in each particular cell factory and of the processes by which the cells interact with the environment imposed on them through the bioreactor operation (Nielsen et al., 2003). The major metabolic pathways of yeasts and the development of the intracellular reaction network of *Pichia pastoris* for metabolic flux analysis is given in detail in Section 2.3.1.1.

In the second step, for each metabolite of the intracellular reaction network one mass balance equation is written. The mass flux balance equations can be represented as the linear vector of differential equations of the form:

$$\mathbf{A} * \mathbf{r}(t) = \mathbf{c}(t) \tag{2.34}$$

where, **A** is the stoichiometric coefficients matrix of the metabolic network, **r**(t) is the vector of fluxes and **c**(t) is the metabolite accumulation vector. The elements of **c**(t) are divided into two subvectors,

$$\mathbf{c}(t) = \mathbf{c}1(t) + \mathbf{c}2(t) \tag{2.35}$$

where, $\mathbf{c}_1(t)$ and $\mathbf{c}_2(t)$ correspond to extracellular and intracellular metabolite accumulation vectors, respectively. Due to the very high turnover of the pools of most metabolites, the concentrations of the different metabolite pools rapidly adjusted to new levels, even after large perturbations in the fermentation broth. Therefore, it may be reasonable to use pseudo-steady state (PSS) approximation for the intracellular metabolites; thus for the intracellular metabolites $\mathbf{c}_2(t)$ can be set to zero. Thus, the obtained set of mass flux balance based stoichiometric linear differential equations forms the mathematical model (Çalık and Özdamar, 2002-a).

The **A** matrix is determined by the biochemistry of the microorganism in the bioreactor, and has a dimension of m x n, where m is the number of metabolites and n is the number of reactions in the metabolic network. If $m > n$, the solution

of the model gives the exact solution. The solution of equation (2.34) can be determined by a constrained least-squares approach with the objective of minimizing the sum of squares of residuals from the stoichiometric mass balance. On the other hand, if *m* < *n,* metabolic flux distributions can be obtained by minimizing or maximizing the objective function, whereupon the best metabolic pathway utilization that would fulfill the stated objective is obtained. In this case, the mathematical formulation for the objective function Z is:

$$Z = \Sigma \alpha_i r_i \tag{2.36}$$

Where Z is a linear combination of the fluxes r_i, and α_i is the coefficient of the component-i in the stoichiometric equation of the corresponding reaction; the metabolic flux distributions are obtained either by minimizing or maximizing the Z in the model. In least-squares method, the matrix is solved and a unique solution is obtained (Çalık and Özdamar, 2002-a).

The only studies in literature, where MFA was performed using *P. pastoris,* is by Sola et al. (2004 and 2007). In the first study, amino acid biosynthesis and central carbon metabolism of *P. pastoris* were studied and compared with those of *Saccharomyces cerevisiae*. It was concluded that, amino acids synthesis and regulation of central carbon metabolism in *P. pastoris* were similar to *S. cerevisiae*. However, in their model, the methanol metabolism had not been included and only glycerol was considered as a substrate, though methanol is important in many cases of recombinant protein production with this organism. In their latter study, the methanol utilization pathway was included in the analysis, however data was obtained from continuous cultures, whereas mostly fed-batch cultures are used, and there was no information about the product.

In the methanol growth phase, methanol is oxidized to formaldehyde by alcohol oxidase at first, and then enters both the dissimilatory pathway and the assimilatory pathway. In the dissimilatory pathway, formaldehyde is oxidized with the generation of NADH to formate and finally to carbon dioxide. In the assimilatory pathway, formaldehyde is phosphorylated to glyceraldehydes-3-phosphate. The subsequent metabolic pathways may again be assumed to similar to those in *S. cerevisiae* (Cereghino and Cregg, 2000). Thus, the

methanol utilization pathway can easily be combined with the central metabolic pathway of *S. cerevisiae*.

Moreover, MFA was applied to other relatives of *Pichia pastoris*; for *Pichia stipitis* (Fiaux et al., 2003) and *Pichia anomala* (Fredlung et al., 2004 a,b), where the central metabolic pathway developed for *S. cerevisiae* was also used.

There are many studies where MFA was applied to *S. cerevisiae* since it is a model organism. Cortassa et al. (1995) and van Gulik and Heijen (1995) constructed metabolic reaction networks for *S. cerevisiae* for the first time. While Cortassa et al. (1995) performed an estimation of the minimal catabolic fluxes needed to fulfill NAD(P) and ATP requirements to sustain growth on various carbon sources and made a comparison of the minimal catabolic fluxes needed with respect to experimental fluxes measured when the yeast grew on different carbon sources, van Gulik and Heijen (1995) estimated the theoretical maximum biomass yields on different substrates, maximum amino acid yields and limit functions to describe the biochemistry of aerobic growth of *S. cerevisiae*. Nissen et al. (1997) reported a stoichiometric model describing for the first time the anaerobic metabolism of *S. cerevisiae* during growth on a defined medium; moreover, this was the first contribution reporting a compartmented model. The fluxes through reactions located around important branch points of the metabolism were compared and the model was used to show the probable existence of a redox shunt across the inner mitochondrial membrane consisting of the reactions catalyzed by the mitochondrial and the cytosolic alcohol dehydrogenase; lastly, they concluded that cytosolic isocitrate dehydrogenase is probably not present during growth on the substrate glucose.

MFA was further applied for *S. cerevisiae* in other studies, to identify transient metabolic regulations (Herwig and von Stockar, 2002), overproduce recombinant proteins (Gonzalez et al., 2003), investigate the effects of bioprocess operation parameters (Wahlbom et al., 2001; Franzen, 2003; Pitkanen et al., 2003; Zhang et al., 2003b), and to characterize the phenotypic response of *S. cerevisiae* to specific genetic modifications (Raghevendran et al., 2004). Furthermore, a comprehensive review of the state-of-the-art for MFA, pertaining to cell growth and synthesis of biomolecules in the commonly used industrial microorganisms was made by Çalık and Özdamar (2002-a). Moreover, a genome-scale metabolic network of *S. cerevisiae* containing 1175 metabolic reactions and 584

metabolites was reconstructed by Förster et al. (2003), providing the first comprehensive metabolic network for an eukaryotic organism.

On the other hand, there is only one study in literature (Çalık and Özdamar, 2002-b) where, MFA was used as a theoretical approach for the analysis of production capacities of erythropoietin by recombinant *Bacillu*s species.

2.3.1.1 Major Metabolic Pathways of Yeasts and Development of the Intracellular Biochemical Reaction Network

A living cell is a complex chemical reactor in which more than 1000 independent enzyme-catalyzed reactions occur. The total of all chemical reactions that occur in the cell is called the metabolism. The metabolic reactions tend to be organized into sequences called metabolic pathways. In this sequence of reactions (the pathway), a precursor is converted into a product through series of metabolic intermediates (metabolites) (Bailey and Ollis 1986). Catabolism is the degradative phase of metabolism, in which organic nutrient molecules (carbohydrates, fats, and proteins) are converted into smaller, simpler end products (e.g., lactic acid CO_2, NH_3). Catabolic pathways release free energy some of which is converted in the formation of ATP and reduced electron carriers (NADH and NADPH). In anabolism, also called biosynthesis, small, simple precursors are built up into larger and more complex molecules, including lipids, polysaccharides, proteins, and nucleic acids. Anabolic reactions require the input of energy, generally in the forms of the free energy of hydrolysis of ATP and the reducing power of NADH and NADPH (Lehninger et al., 1993). Figure 2.7 summarizes the linkage of anabolic and catabolic processes which occur in yeast metabolism. The glycolysis, TCA cycle and oxidative phosphorylation are three central catabolic pathways, and the synthesis of carbohydrates, proteins, lipids and nucleotides are the major anabolic pathways taking place in the microorganism.

Due to the incomplete genome project of *Pichia pastoris* and lack of metabolic studies, the intracellular biochemical reaction network of *Pichia pastoris* was constructed based on the metabolism of the model yeast *Saccharomyces cerevisiae* (Maaheimo et al., 2001; Förster et al., 2003; Kresnowati et al., 2008), the central metabolism of which has also been shown to be suitable for *Pichia stipitis* (Fiaux et al., 2003) and *Pichia pastoris* (Sola et al., 2004 and 2007). The overall pathway was simplified by lumping some reactions into single ones

without loosing representation accuracy. Water, hydrogen, oxygen, sulfate ion and oxidized form of the cofactors (NAD, FAD, NADP) were omitted from the reactions for simplicity (Çalık and Özdamar, 2002-a).

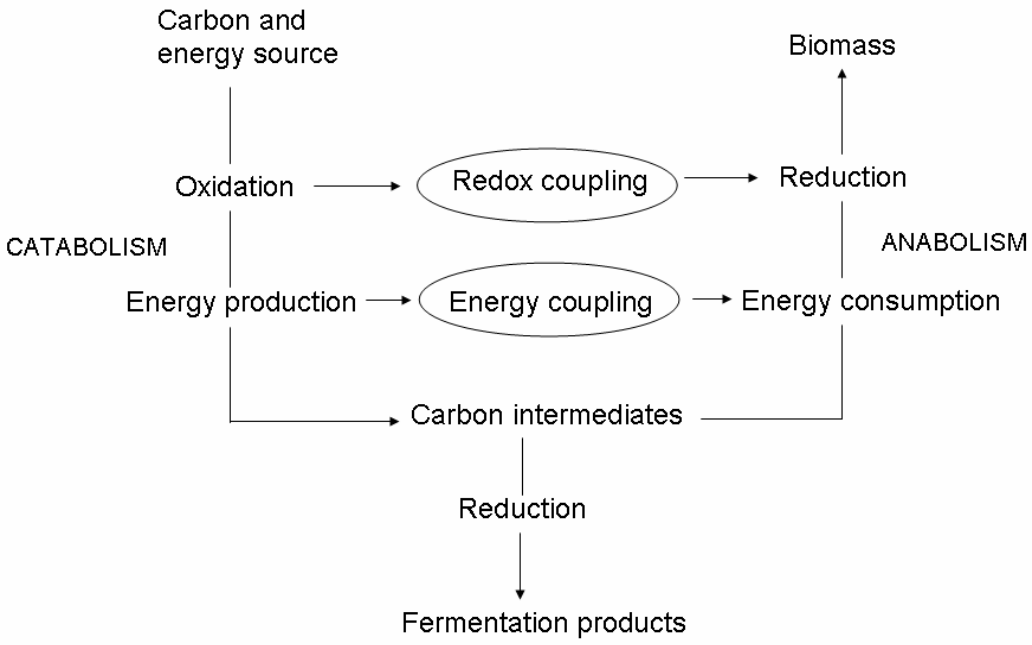


Figure 2.9 Simplistic overview of yeast carbon metabolism (Adapted from Walker et al., 1998)

Since methanol and sorbitol were used as the carbon and energy sources, catabolism reactions of these substrates (Walker, 1998) were included (R1-5) and therefore the reaction for glucose conversion to glucose-6-phosphate was eliminated. The glycolysis and gluconeogenesis pathways are given in R6-15 as separate reactions for ease of calculation. The glycolysis pathway where glucose-6-phosphate is phosphorylated down to pyruvate is regulated to achieve constant ATP levels as well as adequate supplies of glycolytic intermediates that serve biosynthetic roles. The reverse reactions of the glycolysis pathway form the gluconeogenesis pathway. The pentose phosphate pathway (R16-26), which

operates to varying extents in different cells, has two primary functions: (1) to provide cytosolic NADPH for reductive biosynthesis and (2) to provide ribose-5-phosphate for nucleotide and coenzyme biosynthesis (Mathews and van Holde, 1996). In aerobic processes, pyruvate formed in the glycolysis and other metabolites like fatty acids and most amino acids are further oxidized to CO_2 and H_2O via the TCA cycle, which takes place in the mitochondria (R37-47). Therefore, compartmentation into cytosol and mitochondria was considered for metabolites that cannot cross the mitochondrial membrane in both directions (NADH and NADPH) or for acetyl-CoA, which can cross the membrane in only one direction (Förster et al., 2002). Thus, compartmentation of acetyl-CoA, NADPH and NADH between mitochondria and cytoplasm were taken into account by treating these metabolites as if two distinct species in mitochondrial and cytoplasmic reactions. Furthermore, the shuttles for AcCoA, via an acetyl carnitine shuttle, and for NADH, via an ethanol-acetaldehyde shuttle (Bakker 2001), were also included (R31, R126). The replenishment of the intermediates is performed by the anaplerotic reactions given by R34-36 (Förster et al., 2002; Walker, 1998); while the glyoxylate cycle was considered inactive (Maaheimo et al., 2001; Sola et al., 2007).

The biosynthesis of amino acids was specifically shown to be the same as in *Saccharomyces cerevisiae* (Sola et al., 2004). The pathways were grouped as alanine (R51-54), aromatic (R63-66), aspartic (R57-62), glutamic (R67-72) and serine (R48-50) family amino acids, according to their common precursors originating from the intermediary carbon metabolism. Histidine, which has a complex biosynthetic pathway, does not group with the other amino acids (R55-56). With regard to amino acid catabolism (R73-91) by yeasts, several dissimilatory pathways exist including decarboxylation, deamination, transamination and fermentation. Ultimately, the degradation of amino acids and other organic nitrogenous compounds by yeasts yields two end-products, ammonium and glutamate (Large, 1986). Most organisms can synthesize purine and pyrimidine nucleotides from low-molecular weight precursors in sufficient amounts for their needs through the so-called de novo pathways, which are essentially identical throughout the biological world (Mathews and van Holde, 1996). Furthermore, deoxyribonucleotides (dAMP, dGMP, dCMP, dUMP) are synthesized from the corresponding ribonucleotides (AMP, GMP, CMP, UMP) (R90-107). These ribonucleotides and deoxyribonucleotides are the building blocks of RNA and DNA. Single-carbon adducts of tetrahydrofolate (THF) (R111-

115) are utilized in metabolism of serine, glycine, methionine, histidine, among the amino acids, and in purine nucleotides.

The reduced carriers, NADH and $FADH_2$ are reoxidized via the electron transport chain, and the energy released by the transfer of electrons is used to synthesize ATP by the process called oxidative phosphorylation (R116-117). The P/O, which is the number of ATP formed for each oxygen atom used in the oxidative phosphorylation, was taken as 2, which is the theoretical P/O ratio for *S. cerevisiae* (Nielsen et al., 2003). The P/O ratio represents the overall thermodynamic efficiency of the process and differs with the growth conditions. Critical transport reactions were included in R118-126. Carbohydrate biosynthesis (R127) was simplified for yeasts as elsewhere (Nissen et al., 1997), considering carbohydrates as polyglucose (Bruinenberg, 1983).

Lipids in yeast are mainly fats and phospholipids (R128-138), the major building blocks of which are fatty acids (R128-133). The unsaturated fatty acids, which play important roles in yeast physiology (e.g. membrane integrity) include palmitoleic acid (PLLM) and oleic (OLE) acids. Together, these compounds constitute the bulk (approx. 70%) of the fatty acids in *S. cerevisiae* cell membranes. The others consist mainly of the saturated fatty acids: primarily palmitic (PLM) with lesser amounts of stearic acid (STE) (Walker, 1998). The long hydrocarbon chains of fatty acids are efficient ways of energy storage. R133 represents the average fatty acid (FA) formation (Kresnowati et al., 2008).

Biomass (cell) synthesis reaction (R139) is written by considering the macromolecular composition of *S. cerevisiae* (Förster et al., 2003) and assumed to be constant throughout the bioprocess. The synthesis reaction of the recombinant protein EPO (R140) includes the individual amino acids forming the protein, with 4 moles ATP consumed per mole protein-bound amino acid (Mathews and van Holde, 1996). ATP consumption for cellular maintenance is given by R141.

2.3.2 Kinetic Models

As opposed to stoichiometric models, kinetic models describe the dynamic properties of the metabolic network, by combining kinetics with known stoichiometry of metabolic pathways (Gombert and Nielsen, 2000). Kinetic

models in the area of metabolic modeling can be classified as unstructured models and structured models. The details of the reactions occurring within a cell that is considered in the model, determines whether the microbial growth model will be unstructured or structured. Unstructured models simply view the cell as an entity in solution which interacts with the environment; whereas, structured models incorporate the details of intracellular metabolism of the cell. Therefore, structured models are advantageous in describing the dynamic behaviour of the cells. Thus, the structured model obtained may describe the growth process at different operating conditions with the same set of parameters, and it can therefore be applied for optimization of the process or for developing control strategies.

The structured modeling approaches in literature can further be grouped as (Nielsen et al., 2003);

- Simple structured models
  - Compartment models
  - Cybernetic models
- Mechanistic models
  - Single cell models
  - Genetically structured models
- Morphologically structured models.

In compartment models, the biomass is divided into a few compartments or macromolecular pools, by placing cell components with similar function into the same compartment. In cybernetic models, cybernetic variables are introduced into the kinetic model with the aim of substituting the unknown mechanistic details of the cell regulatory architecture by an objective function by supposing that the mechanism of the cell operates with a specific overall goal such as the optimization of growth (Gombert and Nielsen, 2000). Nevertheless, different pathways compete for intracellular resources, which leads to a hierarchically defined optimization criterion with an overall goal (growth) and several sub-goals (pathway resources). From a computational viewpoint, cybernetic modeling poses some difficult problems because a hierarchically defined non-smooth criterion has to be maximized (Wiechert, 2002). In single-cell models, characteristic features of the individual cells are considered. To set up a single

cell model one must have a detailed knowledge of the microorganism, and good single-cell models can therefore be developed only for well examined microorganisms: *E. coli, B. subtilis, and S. cerevisiae* (Nielsen et al., 2003). In morphologically structured models, the differentiation of the cells in the culture is taken into account, which is an important phenomenon for filamentous fungi. However, with yeast cell cultures, describing the average metabolism of the culture using simple structured models gives satisfactory description of many experimental data. Therefore, the mathematical complexity of morphologically structured models can be avoided. Genetically structured models are models for formation of a specific product, rather than cellular growth and can be combined with other models describing the growth of the microorganism in question (e.g. simple structured model), yielding a very detailed process model. The model includes a description at the molecular level of the interactions between the regulatory proteins and genes included in the system of interest. The model thus enables the investigator to examine the variations in the intracellular levels of mRNA and proteins, regulatory as well as structural, at different growth conditions (Agger and Nielsen, 1999). However, gene regulation at the structural level, as well as quantitative data about expression rates and binding affinities are required.

When the model complexity has been specified, mass balances are written for various components of the biomaterial over an appropriately chosen system, and then kinetic rate expressions for the reactions occurring within the biomaterial or at its boundaries are postulated. The combination of mass balances, including kinetic rate equations, and the reactor model makes up a complete mathematical description of the fermentation process. It is then necessary to assign values to the parameters of the model, therefore, the model simulations are compared with the experimental data to estimate the parameter set that gives the best fit of the model to the experimental data (Nielsen et al., 2003).

Only few publications on the unstructured modeling of *P. pastoris* may be found in the literature, despite its increasing importance in the expression of heterologous proteins. First of these was formed by Zhang et al. (2000a) to describe the growth of *P. pastoris* on methanol and optimize the production of a recombinant protein. Jahic et al. (2002) formed a kinetic model that can be used to simulate growth and oxygen consumption and explain two limiting factors for maximum cell density in fed-batch cultures of *P. pastoris*: the maintenance

demand and the feed rate, which is limited by the oxygen transfer capacity. The unstructured model by Pais et al. (2003) describes the relationship between the total biomass and the induction time, both in the batch and fed-batch phases, and correlates recombinant protein production to biomass formation, which also was the main purpose of some other studies (Zhang et al., 2005; Jungo et al., 2006).

On the other hand, there are no structured modeling studies in literature, for *P. pastoris*. Nevertheless, again due to being a model organism, many structured models of various kinds can be found for *S. cerevisiae*, the first being formulated by Bijkerk and Hall (1977). The compartmental structured modeling approach was applied to *S. cerevisiae* for the first time by Villadsen and Nielsen (1990), which was a two compartment model.

Most compartment models are based on a division of the cell into an active and inactive part, where protein synthesizing system is always included in the active part of cell (Nielsen et al., 2003). The inactive part contains rather inactive components such as all of the membrane material. Another approach for dividing the cell into compartments was recently suggested by Aranda et al. (2004) which divides the cell simply into three compartments to predict product accumulation in *S. cerevisiae* cells. The product accumulated in one compartment, the enzymatic compartment was composed of only the enzymes that modify the product, and rest of the components were gathered in the cellular compartment and the experimental data fit reasonable well to the model results.

CHAPTER 3

MATERIALS AND METHODS

3.1 Chemicals

All chemicals and solvents were analytical grade except otherwise stated. The chemicals used in this study and their suppliers are listed in Appendix A.

3.2 Laboratory Equipment

The list of laboratory equipment used in this study is given in Appendix B.

3.3 Buffers and Stock Solutions

All buffers and stock solutions listed in Appendix C were prepared with distilled water. The sterilization of solutions was performed either by autoclaving at 121°C for 20 min or by filter sterilization through 0.20 µm filters (Sartorius AG, Gottingen, Germany).

3.4 Strains, Plasmids and Maintenance of Microorganisms

EPO cDNA was amplified from ORF Clone Collection of Invitrogen (Carlsband, CA, USA). The *Escherichia coli* host strain (clone ID of IOH44362) was carrying the plasmid named pENTRTM221 with a *homo sapiens epo* ORF (NCBI accession number NM_000799) and antibiotic resistance gene to kanamycin. *Pichia pastoris* X-33 wild type strain and pPICZαA shuttle vector carrying Zeocin resistance were purchased from Invitrogen. For all propagations in *E. coli*, One Shot TOP10 chemically component *E. coli* cells (Invitrogen) were used. The *E.*

coli strain carrying the pPICZαA vector was named *E. coli* TOP10-pPICZαA and the strain carrying the vector with EPO gene cloned, was referred to as *E. coli* pPICZαA-EPO. The *P. pastoris* strain obtained at the end of this study, with the EPO gene inserted into its genome was named as *P. pastoris* E17. The strains and plasmids used in this study are summarized in Table 3.1. The sequences and schematic representation of the plasmids are given in Appendix D.

For long-term storage of microorganisms, *E. coli* strains were suspended in 1 mL LB medium containing 15% glycerol and *P. pastoris* strains were suspended in 1 mL YPD medium containing 15% glycerol and both stored at -80°C. The compositions and preparation protocols of all the media used are given in Appendix E.

Table 3.1 Strains and plasmids used in this study

Genus	Species	Strain	Genotype/plasmid	Source
Escherischia	*coli*	TOP10-pPICZaA	pPICZαA	Invitrogen
Escherischia	*coli*	IOH 44362	pENTR221::*EPO*	Invitrogen
Escherischia	*coli*	pPICZaA-EPO	pPICZαA::*epo*	This study
Pichia	*pastoris*	X-33	wild type	Invitrogen
Pichia	*pastoris*	E17	aox1::pPICZaA-EPO	This study

3.5 Growth Media

LB medium containing 50 µg mL^{-1} Kanamycin was used for *E. coli* IOH44362. *E. coli* Top10 cells used for propagation of Zeocin containing plasmids, required the use of SOC medium, and low salt LB (LSLB) medium supplemented with 25 µg mL^{-1} Zeocin for selection. Wild type *P. pastoris* (as the control strain) was grown on YPD agar, transferred to YPD liquid medium and then to BMGY medium for precultivation and finally to BMMY medium for production. For the recombinant strain *P. pastoris* E17, 100 µg mL^{-1} Zeocin was added to the solid and precultivation media, while production medium (BMMY or other defined media) did not require the antibiotic due to genomic integration of the recombinant

gene. The compositions and preparation protocols of all the media used are given in detail in Appendix E.

3.6 Genetic Engineering Techniques

3.6.1 Plasmid DNA Isolation from *E. coli*

Plasmid DNA from *E. coli* cells was extracted using QIAGEN Mini-Prep Kit, (Qiagen, Valencia, CA, USA) according to manufacturer's instructions. For this purpose, a single colony from selective solid medium was inoculated into 5 mL selective liquid medium and grown overnight at 37°C and 200 rpm in 30 mL universal tube. At the end of the protocol, 50 µL of dH_2O was used to elute DNA from QIAprep spin column and the extracted plasmid was stored at -20°C.

3.6.2 Agarose Gel Electrophoresis

Gel electrophoresis was used to separate and visualize DNA strands. 0.8 g of agarose was dissolved in 100 mL 1xTAE and heated until boiling point in a microwave oven. After cooling to approximately 50°C, 0.5 µg mL^{-1} of ethidium bromide was added and the gel was poured into a suitable gel tray. An appropriate comb was inserted and the gel was allowed to cool. The gel tank was filled with 1xTAE buffer. DNA samples mixed with loading buffer were loaded into the wells, together with the DNA ladder (Appendix F) for size estimation. Electrophoresis was performed at 100 V for 45-60 min. The bands were visualized under UV illumination and photographed.

3.6.3 DNA Extraction from Agarose Gels

In the case where a single band of DNA is to be purified from agarose gel after electrophoresis, QIAquick Gel Extraction Kit (Qiagen) was used according to manufacturer's instructions. The procedure is designed to purify DNA from agarose and salts. For this aim, the DNA fragment as visualized above UV illuminating board, was excised from the gel with a clean, sharp scalpel and weighed in an microcentrifuge tube, to be approximately 400 mg of gel. At the end of the protocol, 50 µL of dH_2O was used to elute DNA from QIAquick spin column (Qiagen) and the extracted DNA was stored at -20°C till further use.

3.6.4 Primer Design

The primers used in this study are given in Table 3.2. Primers were designed to amplify the cDNA of *EPO* from the pENTR^{TM}221 plasmid and add *Eco*RI restriction site (6 bp), 6xHis-Tag sequence (18 bp) and factor XA protease recognition sequence (12 bp) to the 5' end of *EPO* sequence during amplification. Since start codon is already present in the vector, it was not included in the primer sequence. Since an addition of 36 bases was required at the 5' end of *EPO* sequence, two relatively short forward primers were designed instead of a single long primer. The reverse primer was designed such that, at the 3' end of the *EPO* sequence, stop codon (taa) and *Xba*I restriction site (6 bp) is present. In the selection of nucleotide sequences corresponding to stop codon and factor Xa recognition sequence (Ile-Glu-Gly-Arg), the codon usage by *P. pastoris* (Appendix G) was also taken into account, wherever possible. Moreover primers for verification of the plasmid sequence were designed (5'AOX1 and 3'AOX1). The sequences for restriction enzyme recognition are underlined. Extra nucleotides were added to the 5' end of restriction enzyme recognition sites, since site recognition by the enzyme close to the end of DNA fragment could otherwise be problematic. The priming sites for sequence verification are given in Appendix D. Formation of primer dimmer and self-complementation are not desired, therefore were controlled by the program Oligo 2.0. The result of this program, together with thermodynamic properties of the designed primers are given in Appendix H.

Table 3.2 Primers used in this study and their sequences

Name	Sequence
EPO-F3-1	5'CACCATATTGAAGGGAGAGCCCCACCACGCCTCATC 3'
EPO-F3-2	5'GGAATTCCACCATCACCATCACCATATTGAAGGGAG 3'
EPO-R3	5'CCACGCTCTAGATTAGTCCCCTGTCCTGC 3'
5'AOX1	5'GACTGGTTCCAATTGACAAGC 3'
3'AOX1	5'GCAAATGGCATTCTGACATCC 3'

3.6.5 Polymerase Chain Reaction (PCR)

The amplification of DNA was carried out by PCR. The reaction mixture, prepared on ice, contained the following:

10x PCR Reaction buffer (with Mg^{2+})	: 5 µL
dNTP (10 mM stock)	: 1 µL
Forward primer 5µM stock	: 3 µL
Reverse primer 5µM stock	: 3 µL
Template DNA	: 100 ng (1-3 µL)
Taq DNA polymerase (1U μL^{-1})	: 1 µL
Sterile dH_2O	: to 50 µL

The PCR program used was:

94°C	3 min	x 1 cycle
94°C	45 sec	x 30 cycles
60°C	45 sec	
72°C	1 min	
72°C	7 min	x 1 cycle
4°C	∞	

3.6.5.1. DNA Purification after PCR

After PCR, removal of primers, nucleotides, polymerases and salts was performed by QIAquick PCR Purification Kit (Qiagen), according to manufacturer's instructions. The purified DNA was eluted with 50 µL of dH_2O from QIAquick spin columns.

3.6.6 Digestion of DNA using Restriction Endonucleases

Restriction enzyme digestion reaction mixture prepared in microcentrifuge tube, contained:

DNA fragment	: 0.1 to 5 µg
10x Reaction buffer (supplied with the enzyme)	: 2 µL
Restriction enzyme	: up to 10 U/µg DNA
Sterile dH_2O	: to 20 µL

The samples were incubated at 37°C water-bath for minimum 2 h. For digestion of genomic DNA, the reaction was carried out overnight. Double digestions using *Eco*RI and *Xba*I could be performed in the same microcentrifuge tube, using a common buffer (Buffer H) recommended by the manufacturer.

3.6.6.1. DNA Purification after Digestion

After restriction endonuclease digestion, removal of primers less than 10 bases, enzymes, salts and unincorporated nucleotides was achieved by QIAquick Nucleotide Removal Kit, according to manufacturer's instructions. At the end of the protocol, purified DNA was eluted with 100 µL of dH_2O from the QIAquick spin column.

3.6.7 Ligation

The 10 µL ligation reaction mixture contained the following:

10X ligation buffer	: 1 µL
Insert DNA (534 bp)	: 45 ng
Double digested vector DNA (3536 bp)	: 100 ng
T4 DNA ligase	: 1 µL
Sterile dH_2O	: to 10 µL

The amount of insert DNA (534 bp) to be added to the reaction mixture was calculated such that insert:vector ratio of 1:3 was achieved, as given in equation 3.1.

$$100 \text{ ng vector} \times \frac{\text{size of insert (bp)}}{\text{size of vector (bp)}} \times \frac{3}{1} = \text{amount of insert (ng)} \qquad (3.1)$$

A control reaction was also set up, containing all the reagents listed above except the insert. The reaction was carried out for 16h in 16°C water-bath equipped with cooler. The ligation product was stored at -20°C till required and 5 µL was used for *E. coli* transformation.

3.6.8 Transformation of *E. coli*

For propagation of plasmids, whether the vector purchased from Invitrogen or the ligation mixture, one-shot *E. coli* TOP10 chemically competent cells were used according to manufacturer's instructions.

One 50 µL vial was thawed on ice for 5 min. 5 µL of the plasmid was added and gently mixed by tapping. The vial was incubated on ice for 30 min, then in 42°C water-bath for exactly 30 sec and immediately placed on ice for 5 min. 250 µL of pre-warmed SOC medium (Appendix E) was added to the vial inside laminar flow cabinet and the vial was incubated at 37°C for exactly 1 hour with shaking. 50-200 µL of transformation mixture was spread on prewarmed LSLB agars + Zeocin (Appendix E) and incubated at 37°C overnight.

3.6.9 DNA Sequencing

The sequence of the *EPO* insert within pPICZαA plasmid was sequenced using primers complementary to the AOX promoter region. 200-500 ng of purified plasmid DNA was mixed with 3 µL of Big Dye reaction mix (version 3.1), 1 µL of 4 µM sequencing primer and the final volume was adjusted to 10 µL with sterile dH_2O.

The PCR block program used was as follows:

96°C	2 min	x 1 cycle
96°C	40 sec	x 30 cycles
50°C	15 sec	
60°C	2 min	
72°C	4 min	x 1 cycle
4°C	∞	

To purify the product, 1 µL of 3 M NaAc (pH 5.2), 1 µL of 125 mM EDTA (pH 8.0), 25 µL EtOH (95%), 1 µL Glycoblue were added into the tube, mixed and left at room temperature for 30 min. The DNA was pelleted (visualised by blue color) by centrifugation at maximum speed for 30min and the supernatant was taken off. 50 µL of 70% EtOH was added and further centrifuged at maximum speed for 15 min. All traces of EtOH were removed by pipeting, the sample was allowed to dry and stored in dark at 4°C, till sequencing.

The sample was run on an ABI 377 fluorescent sequencer by the Sequencing Facility (University of Manchester, UK). The provided data was analyzed using the freeware Chromas program, version 1.45 (http://www.technelysium.com.au/chromas.html, Last accessed: June 2008) and BLAST sequence alignment (http://www.ncbi.nlm.nih.gov/BLAST/, Last accessed: September 2008).

3.6.10 Transfection of *Pichia pastoris*

Transfection of *P. pastoris* was performed using LiCl method according to manufacturer's instructions (Invitrogen, 2002).

YPD plate was inoculated with *Pichia pastoris* X-33 and incubated for 48 h in 30°C incubator. 5 mL YPD was inoculated with a single colony and grown overnight to saturation in 30°C shaker. 50 mL culture of YPD was inoculated using preculture, to an initial OD_{600} of approximately 0.1 and incubated at 30°C with shaking to an OD_{600} of 0.8 to 1.0 (approximately 10^8 cells mL^{-1}; 6-7 h). During this period, the plasmid DNA to be integrated into the genome had to be digested at a single site. Therefore, pPICZαA-EPO plasmid was digested with *Sac*I at 37°C for 5 h in Buffer A. Full digestion was verified by running 2 µL sample on agarose gel and then purified. The concentration of the plasmid DNA was adjusted to 0.1-0.2 µg $µL^{-1}$ using vacuum drying and verified by Nanodrop. When the OD_{600} reached 0.8 - 1.0, the cells were harvested at 4000xg for 5 min, washed with 25 mL of sterile water, and centrifuged at 1500 *g* for 10 min at room temperature. The cell pellet was resuspend in 1 mL of 100 mM LiCl, transferred to a 1.5 mL microcentrifuge tube, centrifuged at maximum speed for 15 sec. LiCl was removed with a pipette and the cells were resuspended in 400 µL of 100 mM LiCl. For each transformation, 50 µL of the cell suspension was dispensed into a 1.5 mL microcentrifuge tube, immediately centrifuged at

maximum speed for 15 sec and LiCl was removed with a pipette. To each tube for transformation, 240 µL of 50% PEG, 36 µL of 1 M LiCl, 10 µL of 5 mg mL^{-1} single-stranded DNA and 5-10 µg plasmid DNA in 50 µL sterile water were added in the order given and vortexed vigorously until the cell pellet is completely mixed. The tube was incubated at 30°C for 30 min without shaking, then heat shocked in a water bath at 42°C for 25 min. The cells were pelleted by centrifugation at 6000 rpm for 15 sec, gently resuspended in 1 mL of YPD and incubated at 30°C with shaking. After 2 h of incubation, 25-100 µL of the medium was spread on YPD + Zeocin plates and incubated for 2-3 days at 30°C.

3.6.11 Isolation of Genomic DNA from Yeast

The isolation of genomic DNA from yeast was performed according to the method described by Burke et al. (2000) with slight modifications.

1. 10 mL yeast culture was grown to saturation in YPD at 30°C in 50 mL Falcon tube.
2. Cells were collected by centrifugation at 5 000 *g* for 6 min.
3. The cells were resuspended in 0.5 mL of dH_2O, transferred to a 1.5 mL microcentrifuge tube and collected by centrifugation at 15 000 *g* for 2 min.
4. The supernatant was decanted and the pellet was vortexed in residual supernatant.
5. 200µl yeast lysis solution was added and mixed by inversion to ensure lysis of cells and then 200µl phenol:chloroform:isoamyl alcohol (25:24:1) and 0.3 g acid-washed glass beads were added.
6. The tube was wrapped with Parafilm and vortexed for 3-4 min.
7. 0.2 mL of TE (pH 8.0) buffer was added and centrifuged at 15 000 *g* for 5 min in a microcentrifuge.
8. The aqueous layer was transferred to a fresh microcentrifuge tube.
9. 1 mL of 100% ethanol was added and mixed by inversion.
10. The tube was centrifuged at 15000 *g* for 2 min and the supernatant was discarded.
11. The pellet was resuspended in 0.4 mL of TE buffer and 3 µL of a 10 mg mL^{-1} solution of RNase A.

12. The solution was incubated for 10 min at 37 °C and 10 µL of 4 M ammonium acetate plus 1 mL of 100% EtOH were added and mixed by inversion.
13. The DNA was pelleted by centrifugation at 15 000xg for 5 min in a microcentrifuge and the supernatant was discarded. The pellet was air-dried and resuspended in 50 µL of sterile dH_2O.
14. 10 µL was used for each sample to be analyzed by Southern blotting. This corresponds to approximately 2-4 µg of genomic DNA.

3.6.12 Southern Blotting

Southern blotting was performed to confirm genomic integration of pPICZαA-EPO into *P. pastoris* X-33 genome. The detailed protocol is given in Appendix I. Shortly, the isolated yeast genomic DNA (Section 3.6.11) was digested with *Bam*HI at 37°C overnight, purified with EtOH precipitation and run on 200 mL, 0.7% agarose gel, for 5h at 60 V. VacuGene XL vacuum blotting system (Amersham) was used for blotting. The system was operated under 50 mbar vacuum while the gel was covered with depurination, denaturation and neutralising solutions (Appendix C) for 20 min each and then with transfer solution for 3 h; removing the solutions completely after each step. After blotting, the membrane was dried between blotting papers in 80°C oven for 2 h.

In the mean time, the probe was labelled using AlkPhos Direct Labeling (Amersham) according to manufacturer's instructions. EPO gene was amplified by PCR (Section 3.6.5) using the primers EPO-F3-1 and EPO-R3, as the probe to be labeled.

The membranes were prehybridized for 15 min in 55°C hybridization oven using 50 mL hybridization buffer, then 500 ng of the labeled probe was added and hybridization was carried out overnight. Post hybridization stringency washes were performed again in 55°C hybridization oven, using primary and secondary wash buffers in excess amount, 20 min for each buffer, changing to fresh buffer after 10 min. For chemiluminescent detection, 7 mL of CDP-Star detection reagent was poured on the membrane and left for 5 min. The chemiluminescent signals were detected by exposing to X-ray film.

3.7 Recombinant Protein Production and Purification

3.7.1 Recombinant Protein Production

Protein production was carried out either in batch cultures using laboratory scale air filtered shake bioreactors or pilot scale bioreactors.

3.7.1.1 Precultivation

RHuEPO producing *P. pastoris* E17 strain from glycerol stocks was streaked onto YPD solid medium containing 0.1 g L^{-1} Zeocin, and incubated for 48 h at 30°C then a single colony was inoculated into 5 mL YPD + 100 µg mL^{-1} Zeocin medium and shaken at 225 rpm, 30°C overnight. Thereafter, the BMGY precultivation medium was inoculated from this seed culture and cells were grown for 20 - 24 h to reach OD_{600} 2-6. The cells were harvested by centrifugation at 2000xg for 5min and resuspended in production medium (in laboratory scale or pilot scale bioreactors) such that the initial OD_{600} =1 was obtained.

3.7.1.2 RHuEPO Production in Laboratory Scale Air Filtered Shake Bioreactors

Recombinant protein production in batch cultures were carried out in baffled and air filtered V=250 mL Erlenmeyer flasks containing V_R=50 mL production medium (either BMMY or BSM+PTM1 given in Appendix E), for up to 72 h, and every 24 h, methanol was added to the medium to 1% (v/v).

3.7.1.3 RHuEPO Production in the Pilot Scale Bioreactor

The pilot scale 3.0 dm^3 bioreactor (Braun CT2-2), having a working volume of 0.5-2.0 dm^3 and consisting of temperature, pH, foam, stirring rate, feed inlet rate and dissolved oxygen control was used. The bioreactor utilized an external cooler, steam generator and a jacket around the bioreactor for sterilization and temperature control. The bioreactor was stirred with two four-blade Rushton turbines and consisted of four baffles and a sparger. Air was supplied through a compressor and oxygen through a pure oxygen tube. The inlet air was enriched with oxygen passing thorough a mass flow controller when needed. Feed solutions placed on balances were transferred aseptically through inlet ports using peristaltic pumps.

A rather standard protocol for expression of recombinant proteins in *P. pastoris* under the control of the *AOX1* promoter was followed (Stratton et al., 1998) with slight modifications, for production at high cell densities in controlled bioreactors. The cells, harvested from the precultivation medium and resuspended in the defined medium inside the bioreactor, were first grown in batch mode until glycerol is totally consumed, in order to achieve high cell densities rapidly. Secondly, glycerol feed (50% glycerol containing 12 mL L^{-1} PTM1) was fed at limiting concentrations by a predetermined rate in order to increase biomass concentration further, while derepressing the AOX enzyme (necessary for the dissimilation of methanol) gradually and reducing the time necessary for the cells to adapt to growth on methanol. A short transition phase of methanol was performed by giving a short pulse of methanol feed, C_{M0}=1.5 g L^{-1}, (from 100% methanol containing 12 mL L^{-1} PTM1) and after 3 hours, recombinant protein production was induced by the addition of methanol in fed-batch mode, again with a predetermined feeding profile. In experiments containing sorbitol, sorbitol (450 g L^{-1} sterile solution) was added batch-wise to the system at the end of methanol transition phase, such that C_{S0}=50 g L^{-1}.

The steps of scale up and the bioreactor system are shown in Figure 3.1. At the end of the production, the medium was centrifuged at 13000 *g* for 10 min, the cells were discarded and supernatant was stored at 4°C till ultrafiltration or at -20°C for further analysis.

3.7.1.4 Recombinant Protein Production using Biodiesel Byproduct Crude Glycerol

RHuEPO-producing *P. pastoris* E17 strain from the seed culture was inoculated into buffered complex medium containing (g L^{-1}): yeast extract, 1; peptone, 2; YNB, 1.34; biotin, $4x10^{-5}$; glycerol, 12.6; and 0.1 M potassium phosphate buffer (pH 6.0). Media differed by the source of glycerol used, which was either pure commercial glycerol (Merck) or crude glycerol, the byproduct of biodiesel production. Crude glycerol samples were synthesized using different vegetable oils by studies in Gazi University (Çelik et al., 2008). Cell growth was allowed to proceed until glycerol was totally consumed (at about t=30 h), as detected by immediate off-line HPLC analysis. At this point (t=30 h), methanol was added to all media as the inducer of recombinant protein production, at the final concentration of 8 g L^{-1} (1% v/v). Although the medium initially contained

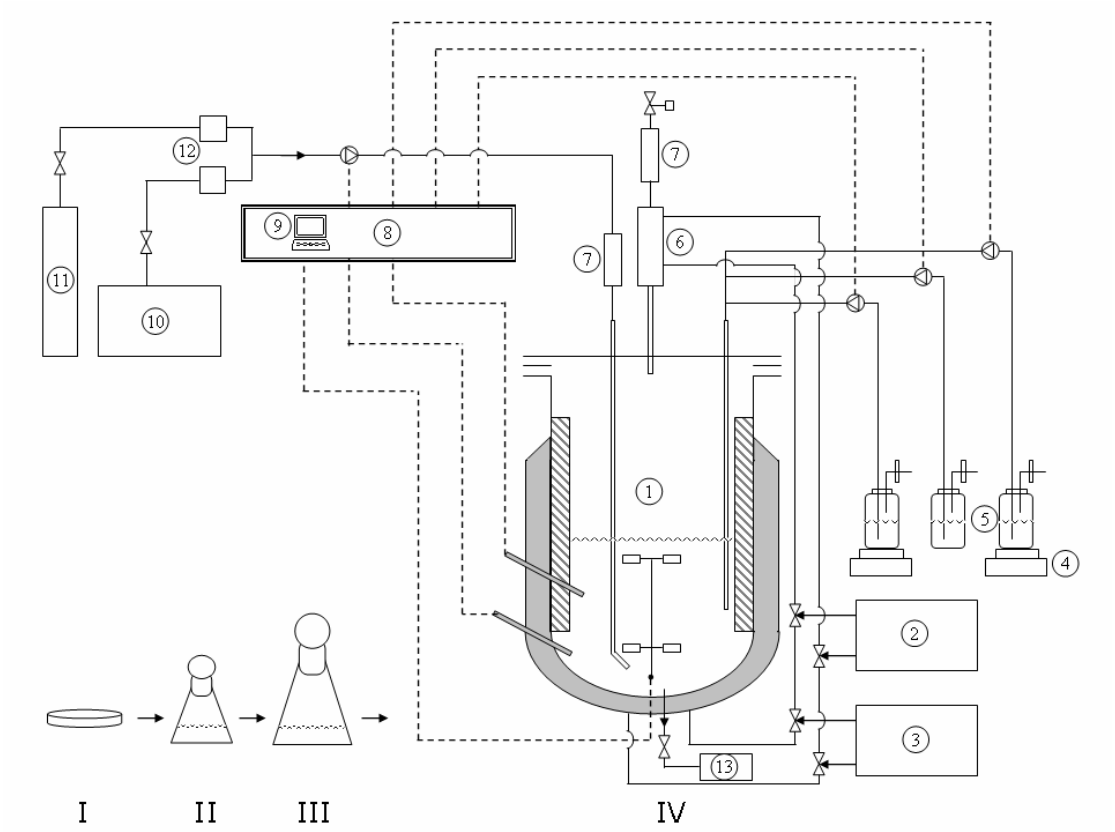


Figure 3.1 Scale up steps and the pilot scale bioreactor system. I: Solid medium inoculated from stock culture; II: 1st Precultivation medium, V = 10 cm^3; III: 2nd Precultivation medium, V = 100 cm^3; IV: Pilot scale bioreactor system, Vo=800 cm^3, which is composed of (1) Bioreaction vessel, Biostat CT2-2 (2) Cooling circulator (3) Steam generator (4) Balances (5) Feed, base and antifoam bottles (6) Exhaust cooler (7) Gas filters (8) Controller (9) Biostat CT Software (10) Air compressor (11) Pure O2 tank (12) Digital mass flow controllers (13) Sampling bottle.

methanol as well, the toxicity level of methanol (>3%) was not reached in any medium, at any time of the bioprocess. Protein production was carried out in duplicate batch cultures, using baffled 250 mL Erlenmeyer flasks containing 50 mL of production medium. The samples taken were centrifuged at 13 000 *g* for 10 min, the cell pellet was discarded, and the supernatant was used in further analysis.

3.7.2 Ultrafiltration

Concentration and desalting of production medium (supernatant) was achieved by ultrafiltration using 400 mL stirred cells (Amicon) with 10 kDa cut-off regenerated cellulose ultrafiltration membranes (Millipore). The process was carried in cold room (2-8°C) using N_2 gas pressure of maximum 55 psi (3.8 bar), until at least 10-fold concentration of the medium was obtained.

3.7.3 Polyhistidine-Tag Affinity Purification of rHuEPO

The recombinant protein, rHuEPO was purified using its polyhistidine tag, with cobalt-based metal affinity resins (BD Talon). 1 mL of resin was used for 20 mL production medium after ultrafiltration and incubated in cold room overnight on rotary platform. Resin bound polyhistidine tagged recombinant proteins were eluted with imidazole containing elution buffer. The detailed protocol is given in Appendix J.

3.7.4 Factor Xa Digestion of rHuEPO

After purification of the polyhistidine tagged recombinant protein, the 6xHis Tag was digested using the factor Xa protease recognition sequence cloned in the 5' end of the EPO sequence, to obtain the native protein when needed. However, since factor Xa protease is sensitive to imidazole used in elution step of affinity purification, the sample was first desalted using ultrafiltration spin columns with 10 kDa cut-off (Vivaspin 500, Vivascience, Germany). 10 µg protein sample was digested using 1 U of factor Xa protease (Qiagen) at 25°C for 16 h. The protease was removed using Xa removal resin provided with the enzyme, according to manufacturer's instructions. The native protein was further purified from the cleaved 6xHis peptides and the undigested 6xHis-tagged protein using the cobalt-based metal affinity resins as described in Section 3.7.3.

3.7.5 Deglycosylation of rHuEPO

For digestion of *N*-glycans, 50 - 500 µg of glycoprotein rHuEPO was dissolved in 45 µL of reaction buffer (20 m*M* sodium phosphate; 0.02% sodium azide, pH 7.5). 2.5 µL of denaturation solution (2% SDS, 1 M β-mercaptoethanol) was added and the glycoprotein denatured by heating at 100°C for 5 min. After cooling to room temperature, 2.5 µL of detergent solution (15% NP-40) and 2 µL of N-Glycanase were added to the reaction mixture and incubated overnight at 37°C. For removal of *O*-glycans (optional), the product of the previous reaction was incubated in the presence of 0.02 M NaOH at 50°C for 24 h.

3.8 Analysis

3.8.1 Cell Concentration

Cell concentration was measured using a UV-Vis Spectrophotometer at 600 nm and 1 mL sample taken from liquid medium was diluted with dH_2O to read OD_{600} in the range 0.1 - 0.8. The calibration equation used for converting absorbance to cell concentration, C_X (kg m^{-3}) is given in equation (3.2).

$$C_X = 0.3 * OD_{600} * \text{Dilution Ratio} \quad (3.2)$$

3.8.2 Total Protein Concentration

Total protein concentration was determined spectrophotometrically using Bradford assay (Bradford, 1976). 20 µL of sample was mixed with 1 mL of Bradford reagent (BioRad), incubated at room temperature for 5 min and the absorbance was read at 595 nm by UV-spectrophotometer. The calibration curve was obtained using BSA in the concentration range of 0-2 mg mL^{-1} (Appendix K).

3.8.3 RHuEPO Concentration

RHuEPO concentrations were measured using a high-performance capillary electrophoresis (Waters HPCE, Quanta 4000E, Milford, MA). Samples filtered with 0.45 µm cellulose acetate filters, which provide least protein adsorption, were analyzed at 12 kV and 15°C with a positive power supply using 60 cm x 75 µm silica capillary and 50 mM borate buffer (pH=10) containing zwitter ion (Z1-Methyl reagent, Waters) as the separation buffer. The zwitter ion is used to

prevent protein adsorption to the capillary column. Proteins were detected by UV absorbance at 214 nm, as mentioned elsewhere (Çalık et al., 1998). RHuEPO purified by ultrafiltration followed by polyhistidine-tag affinity purification (Çelik et al., 2007) was used as the standard, whose concentration was determined by the Bradford assay (Bradford, 1976).

The rHuEPO peak was detected at a time ratio of 1.782, with respect to the reference sodium peak. Thus, if the sodium peak is at 11.9 min for example, the rHuEPO peak should be expected around 21.2 min. The calibration obtained from the spectra was 0.002 g rHuEPO area^{-1} m^{-3}, using 10 µL sample volume.

3.8.4 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE was performed as described by Laemmli (1970). Purified protein samples were mixed with the sample buffer and heated in boiling water for 4 min. 20 µL of the samples and 5 µL of a dual color prestained protein MW marker (Appendix F) was run simultaneously at 40 mA of constant current. The buffers used are given in Appendix C and the detailed protocol is given in Appendix L.

3.8.4.1 Staining the SDS-PAGE Gels

The SDS-PAGE gels that were not going to be used in western blotting were stained using either coomassie stain or silver stain and photographed. For coomassie staining, the gel was washed twice with water, and shaken 1-4 h at room temperature in coomassie blue stain solution (Appendix C) covering the gel. After washing with water, the gel was shaken in destain solution (Appendix C) for at least 2 h before the bands become visible. The gels were silver stained using the procedure of Blum *et* al. (1987), the detailed protocol is given in Appedix M.

3.8.5 Western Blotting

Western Blotting was performed to distinguish EPO by using its specific antibody. After the samples were run on SDS-PAGE gel, the gel was sandwiched between membrane, blotting paper, sponges and the blotting cassette. The membrane (PVDF) used, had been pre-wet with methanol and equilibrated with the Transfer Buffer. Blotting was performed electrophoretically, in the cold room for 3 h at 50 V. The blotted membrane was washed with TBS-T 2-3 times, then with TBS-T-

Milk for 1 h, at room temperature, shaking. After 3 washes with TBS-T, the membrane was incubated overnight at 4°C, in primary antibody diluted 1:1000 with TBS-T-Milk. After 3 washes with TBS-T, the membrane was incubated for 1 h at room temperature, in secondary antibody diluted 1:10000 with TBS-T-Milk. After 3 washes, the presence of rHuEPO was visualized either by peroxidase-based immunocytochemical procedure using a diaminobenzidine chromogen kit (Biomeda Corp., Foster City, CA) or using a chemiluminescent substrate (Pierce), the signals of which were detected by exposing to X-ray film. Detailed protocol is given in Appendix L.

3.8.6 Mass Spectrometry

Protein identification was made by producing a tryptic digest of protein corresponding to rHuEPO excised from Coomassie Brilliant Blue G-250-stained SDS-PAGE gel, and then determining peptide mass fingerprints on a quadrupole time-of-flight mass spectrometer (Q-ToF MS) using electrospray ionization (Waters-Micromass, Milford, MA).

The molecular weights of the glycans, of intact rHuEPO and of deglycosylated rHuEPO were analyzed by the use of a MALDI-LR (Waters-Micromass, UK) instrument. All spectra were generated using a pulsed nitrogen gas laser (337 nm) in positive-ion mode with built-in delayed extraction. Accelerating voltage was 15 kV. Deglycosylated protein was precipitated by the addition of ice-cold ethanol (Küster et al., 1997). The glycans, dissolved in the supernatant, were recovered and dried by centrifugal evaporator. All samples were desalted by drop dialysis prior to analysis (Gorisch, 1988). 3 µl of 10 mg mL^{-1} matrix dissolved in 50% acetonitrile and 0.1% TFA solution, was mixed with 1 µl of approximately 10 pmol μl^{-1} sample and 1 µl of this mixture was spotted on target plate and air dried ("dried droplet" technique; Karas and Hillenkamp, 1988). 0.5 µl of ethanol was used for recrystallization on the plate, in case of glycan samples. Sinapinic acid matrix was used for protein samples and Super-DHB (Sigma, St. Louis, MO) matrix for glycan samples. For detection of proteins, linear mode with a low mass gate of 1000 Da and for glycans, reflectron mode with a low mass gate of 800 Da was used. $Man_9(GlcNAc)_2$ (Sigma, St. Louis, MO) oligosaccharide and intact protein cytochrome c were used as external molecular weight standards. Spectra were generated from the sum of 100-200 laser pulses

and mass determinations were made by finding the peak centroid of a smoothed signal (by Savitzky-Golay algorithm) after background subtraction.

3.8.7 Glycerol, Methanol and Sorbitol Concentrations

Glycerol, methanol and sorbitol concentrations were measured by reversed phase HPLC (Waters HPLC, Alliance 2695, Milford, MA) on Capital Optimal ODS-5μm column (Capital HPLC, Scotland) at 30°C, using 5 mM H_2SO_4 as the mobile phase at a flow rate of 0.5 mL min^{-1}, and detected with refractive index detector (Waters-2414) at 30°C. The method is based on reversed phase HPLC, in which the concentrations were calculated from the chromatogram, based on the chromatogram of the standard solutions. Samples were filtered with 0.45 μm cellulose acetate filters and loaded to the analysis system. The analysis was performed under the optimized conditions specified below:

Column	: Capital Optimal ODS, 5μm
Column dimensions	: 4.6×250 mm
System	: Reversed phase chromatography
Mobile phase and flow rate	: 5 mM H_2SO_4, 0.5 mL min^{-1}
Column temperature	: 30°C
Detector and temperature	: Waters 2414 Refractive Index detector, 30°C
Injection volume	: 5 μL
Analysis period	: 10 min

The calibration curves are given in Appendix N.

3.8.8 Amino Acids Concentrations

Amino acid concentrations were measured with an amino acid analysis system (Waters, HPLC), using the Pico Tag method (Cohen, 1983). The method is based on reversed phase HPLC, using a precolumn derivation technique with a gradient program developed for amino acids. The amino acid concentrations were calculated from the chromatogram, based on the chromatogram of the standard amino acids solution. The analysis was performed under the conditions specified below:

Column	:Amino acid analysis (Nova-Pak C18, Millipore)
Column dimensions	:3.9 mm x 30 cm
System	:Reversed phase chromatography
Mobile phase flow rate	:1 mL min^{-1}
Column temperature	:38 °C
Detector and wavelength	:UV/VIS, 254 nm
Injection volume	:4 µL
Analysis period	:20 min

3.8.9 Organic Acids Concentrations

Organic acid concentration were measured with an organic acid analysis system (Waters, HPLC, Alliance 2695). The method is based on reversed phase HPLC, in which organic acid concentrations were calculated from the chromatogram, based on the chromatogram of the standard organic acids solution. Samples were filtered with 0.45 µm filters and loaded to the analysis system with a mobile phase of 3.12% (w/v) NaH_2PO_4 and 0.62×10^{-3}% (v/v) H_3PO_4 (İleri and Çalık, 2006). The analysis was performed under the conditions specified below:

Column	: Capital Optimal ODS, 5µm
Column dimensions	: 4.6×250 mm
System	: Reversed phase chromatography
Mobile phase flow rate	: 0.8 mL min^{-1}
Column temperature	: 30°C
	: Waters 2487 Dual absorbance detector,
Detector and wavelength	: 254 nm
Injection volume	: 5 µL
Analysis period	: 15 min

3.8.10 Protease Activity Assay

Proteolytic activity was measured by hydrolysis of casein. The culture broth was harvested by centrifugation at 13500 *g* for 10 min. Hammerstein casein (2 mL of

0.5% w/v) in either 0.05 M borate buffer (pH 10), 0.05 M sodium acetate buffer (pH 5) or 0.05 M sodium phosphate buffer (pH 7) was mixed with 1 mL of diluted sample and hydrolyzed under T=37°C, pH=10 for 20 min. The reaction was stopped by adding 10% (w/v) trichloroacetic acid (TCA) and the reaction mixture was centrifuged at 10500 *g* for 10 min at +4°C, then the absorbance of the supernatant was measured at 275 nm in UV-Vis spectrophotometer. One unit protease activity was defined as the activity that liberates 4 nmole tyrosine min^{-1}. (Moon and Parulekar, 1991). The calibration equation used for converting absorbance to protease activity (U cm^{-3}) (Çalık, 1998) is given in equation (3.3).

$$A = \left(\frac{Absorbance}{0.8x1/\mu mol.cm^{-3}}\right)\left(\frac{1U}{4nmol/\min}\right)\left(\frac{1}{20\min}\right)\left(\frac{1000nmol}{1\mu mol}\right)\left(\begin{matrix}Dilution\\Ratio\end{matrix}\right) \qquad (3.3)$$

3.8.11 AOX Activity Assay

A bi-enzymatic assay comprising alcohol oxidase (AOX) and horseradish peroxidase (HRP) was used to monitor the oxidation of methanol to formaldehyde by AOX. A colorimetric system based on the combination of phenol-4-sulfonic acid (PSA) and 4-aminoantipyrine (4-AAP) was chosen to measure the concentration of H_2O_2 produced by AOX. In this particular system, two moles of H_2O_2 react with one mole of PSA and one mole of 4-AAP, yielding three moles of water, one mole of sodium hydrogenosulfate and one mole of a quinoneimine dye.

Methanol + O_2 $\xrightarrow{AOX}$ Formaldehyde +H_2O_2

2 H_2O_2 + PSA + 4-AAP $\xrightarrow{HRP}$ Quinoneimine dye + 3 H_2O + $NaHNO_3$

This dye has a characteristic magenta color with maximum absorption around 490 nm. The activity of AOX was determined by monitoring the associated increase in absorbance at 490 nm with UV-Vis spectrophotometer. This increase is proportional to the rate of H_2O_2 production and, consequently, to the rate of methanol consumption. All kinetic studies were performed at 25°C using a standard assay reaction mixture, containing 0.4mM 4-AAP, 25mM PSA, and 2 U

mL^{-1} HRP in 0.1M phosphate buffer, pH 7.0. One unit of activity (U) was defined as the number of μmol of H_2O_2 produced per minute at 25°C (Azevedo et al., 2004).

3.8.12 Liquid Phase Mass Transfer Coefficient and Oxygen Uptake Rate

The Dynamic Method (Bandyopadhyay and Humprey, 1967), as explained in Section 2.2.2, was used in order to determine the liquid phase mass transfer coefficient and oxygen uptake rate in the rHuEPO production process.

Prior to inoculation of the microorganism to the production medium in the bioreactor, the physical mass transfer coefficient (K_La_0) was determined. After inoculation of the microorganism to the bioreactor, the dynamic oxygen transfer experiments were carried out at certain cultivation times in a short period of time, so that the microorganisms are unaffected.

3.9 Mathematical Modeling

3.9.1 The Stoichiometric Model and Mass Flux Balance-Based Analysis

The central metabolic reaction network of *Saccharomyces cerevisiae* (Maaheimo et al., 2001; Förster et al., 2003; Kresnowati et al., 2008) was used to form the central metabolic reaction network of *P. pastoris*. The reaction network formed for *P. pastoris* given in Appendix O sums up to a total of 102 metabolites and 141 reactions. The optimization program GAMS 2.25 (General Algebraic Modeling System, GAMS Development Corp., Washington DC) was used to solve mass-flux balance equation (2.34) given in Section 2.3.1. Using mathematical programming, optimum flux distributions were obtained by minimizing the objective function Z (equation 2.36). The model variables that are the reaction fluxes of metabolites were expressed in mmol g^{-1} DW h^{-1}; and the flux towards biomass represents the specific growth rate, μ in h^{-1} (Çalık et al., 1999).

3.9.2 The Kinetic Model

The structural, or physiological, origins of the model was defined first, thus the biomass was divided into three compartments, or macromolecular pools, according to the physiology of the yeast and by placing cell components with similar function into the same compartment. Thus, the recombinant product was

considered in one compartment, alcohol oxidase and protease enzymes were gathered in the enzymatic compartment and rest of the cell components were gathered in the cellular compartment. Thereafter material balances were written for each component, and the kinetic rate equations were written, again based on physiology of the yeast. The kinetic constants were determined based on experimental data or from literature. The differential equation set was then solved using Polymath, to verify that the model fits experimental data.

CHAPTER 4

RESULTS AND DISCUSSION

In this study, an expression system for production of the therapeutically important glycoprotein, recombinant human erythropoietin (rHuEPO), by *P. pastoris* was designed and investigated. In the first research phase of the study, the recombinant microorganism producing rHuEPO was developed and the recombinant protein produced by the microorganism was purified and characterized by Western blotting and MALDI-ToF MS. In the second research phase of the study, the effects of alternative fermentation medium components and pH on the expression of rHuEPO and cell growth were analyzed in laboratory scale air filtered shake bioreactors. Using the designed defined medium, the effects of feeding strategy on cell growth, rHuEPO production, byproduct formation, oxygen transfer and fermentation characteristics were investigated by using pilot-scale fed-batch bioreactors. Moreover, a stoichiometric and a kinetic model was formed, with the aim of better describing the cell growth and recombinant protein production in *P. pastoris* mathematically.

4.1 Development of the Recombinant *Pichia pastoris* Producing RHuEPO

The research program, schematically summarized in Figure 4.1, for the development of r-*P. pastoris* strain producing rHuEPO was carried out mainly in three parts. Firstly, the EPO gene was fused with a polyhistidine tag and a target site for the factor Xa protease by primer extension and PCR amplification approaches. Thereafter, the fused gene was cloned into vector pPICZαA, which carries the α-factor signal peptide, *AOX1* promoter and Zeocin resistance gene. Finally, the recombinant vector was introduced and expressed in *P. pastoris*.

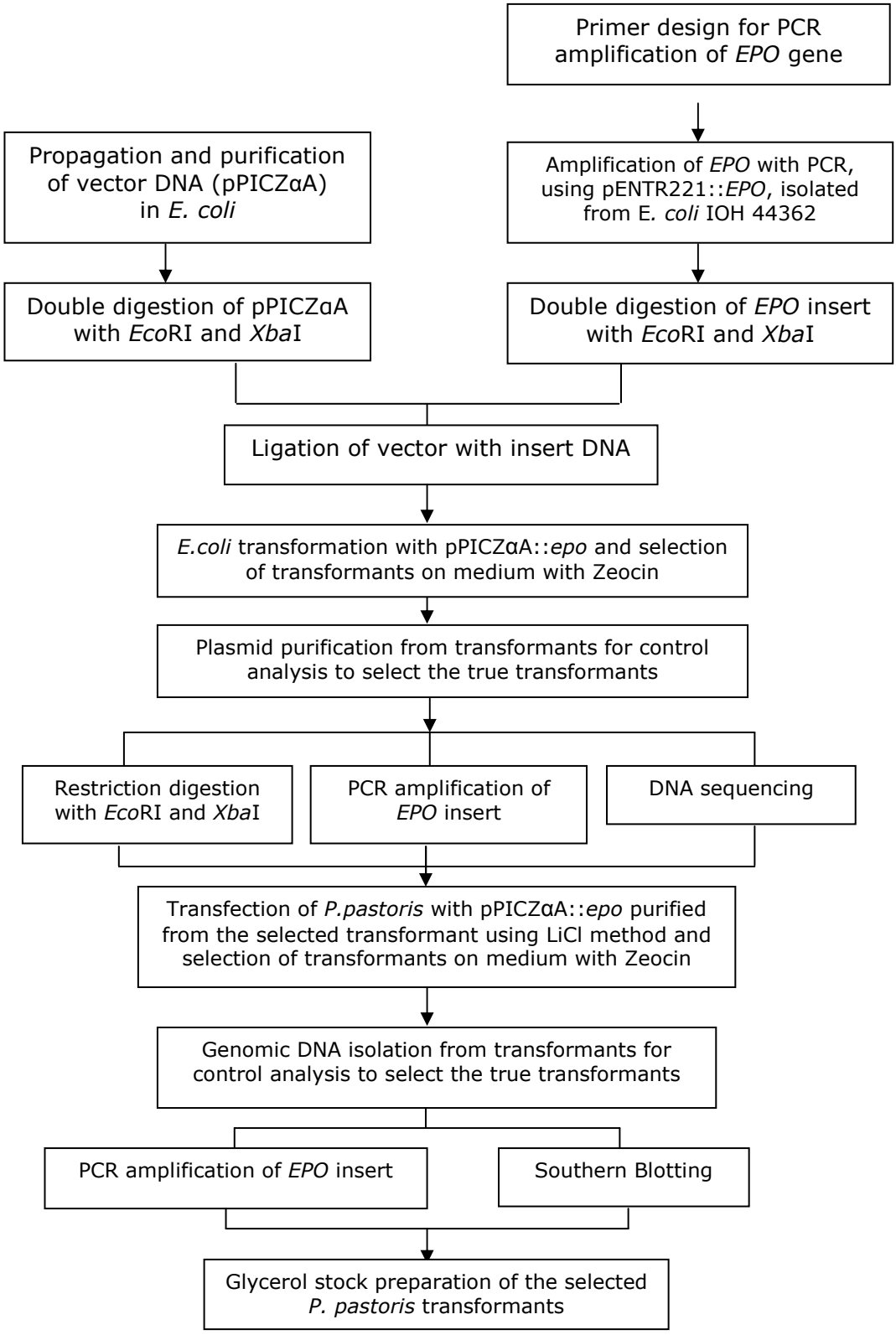


Figure 4.1 Flowchart of the research plan, for the development of the recombinant *Pichia pastoris* producing extracellular rHuEPO.

4.1.1 Propagation and Purification of pPICZαA

The pPICZαA vector delivered as a lyophilized powder, was propagated in *E. coli* Top10 cells. The powder was dissolved in 20 µL of sterile dH_2O, and 4 µL of it was used in transformation (Section 3.6.8). From a single colony, glycerol stocks were prepared (Section 3.4). The plasmid was purified from this colony (Section 3.6.1) and digested once to visualize its actual size (3593 bp) in agarose gel electrophoresis (Lane 1, Figure 4.2). Undigested plasmids run faster on agarose gel and are not useful in visualizing the actual size of the plasmid (Lane 2, Figure 4.2).

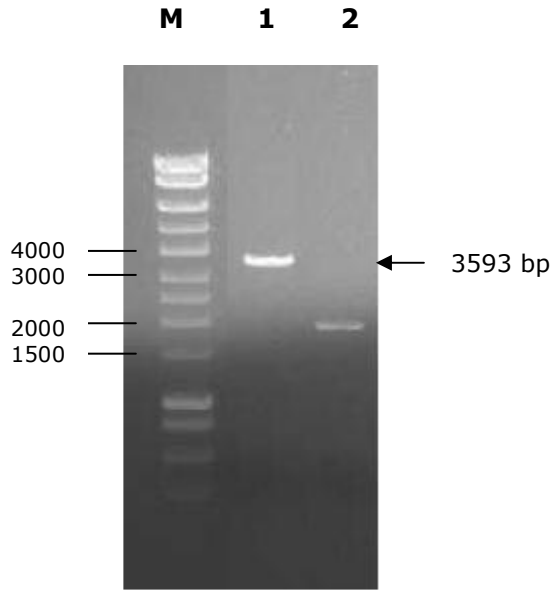


Figure 4.2 Agarose gel electrophoresis of pPICZαA extracted from *E. coli* Top10. M: Hyperladder I, Lane 1: pPICZαA-single digested, Lane 2: undigested pPICZαA.

4.1.2 Amplification of EPO cDNA from *E. coli* IOH44362 by PCR

E. coli IOH44362 cells from glycerol stocks initially prepared were plated onto LB plates containing 50 µg mL^{-1} Kanamycin and then inoculated into LB liquid media containing 50 µg mL^{-1} Kanamycin for plasmid extraction (Figure 4.3.a). 2 µL of the extracted plasmid was used in PCR amplification of EPO gene (Section 3.6.5), using the primer set EPO-F3-1 and EPO-R3. The purified PCR product was run on agarose gel (Figure 4.3.b), and the 528 bp bands were extracted from the gel (Section 3.6.3), to be used as template in the second set of PCR.

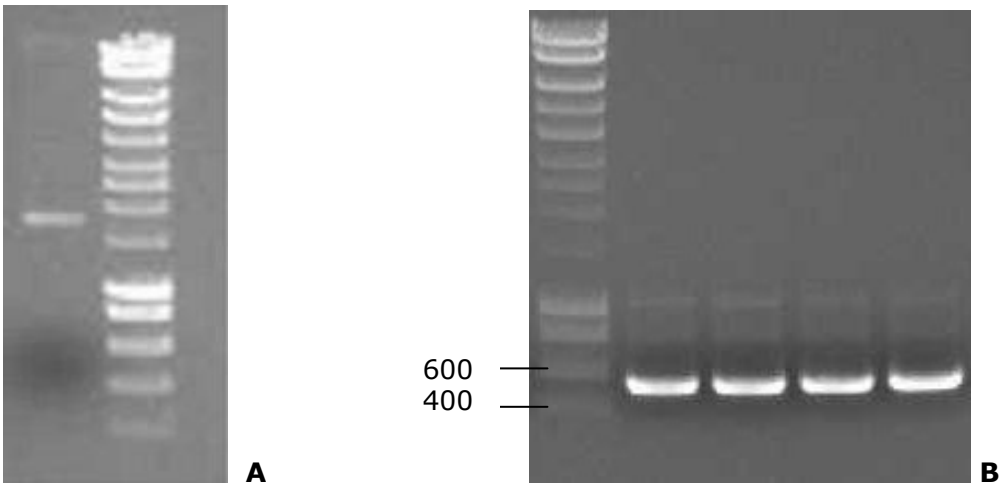


Figure 4.3 Agarose gel electrophoresis results A. Plasmid (pENTR221::EPO) extracted from *E. coli* IOH44362. B. EPO gene amplified by PCR, using pENTR221::EPO as template and the primer set EPO-F3-1 and EPO-R3. The 528 bp product was purified from this gel.

In the second set of PCR, the 3 µL of the product purified from the previous PCR was used as the template with the primer set EPO-F3-2 and EPO-R3. The aim of this reaction was to elongate the 5' end of EPO gene, as schematically explained in Figure 4.4. The C-terminal polyhistidine tag already present in the pPICZαA vector was not employed because its removal using known methods would not achieve the goal of producing a mature recombinant protein that was the native form of EPO. For this reason, a factor Xa recognition site was placed immediately at the *N*-terminal end of EPO, which enabled use of this protease to digest the non-native *N*-terminal region of EPO after its secretion and affinity purification, as will be explained later in detail.

4.1.3 Digestion and Ligation Reactions

The restriction digestion enzymes used in this study were selected with the following criteria:

- *Eco*RI and *Xba*I are at either ends of the multiple cloning site of the vector pPICZαA, so there would not be any extra nucleotides in the sequence.
- They have no restriction sites in the sequence of EPO gene as verified by NEBcutter V2.0 software (http://tools.neb.com/NEBcutter2/index.php).
- They are relatively inexpensive and readily available enzymes.

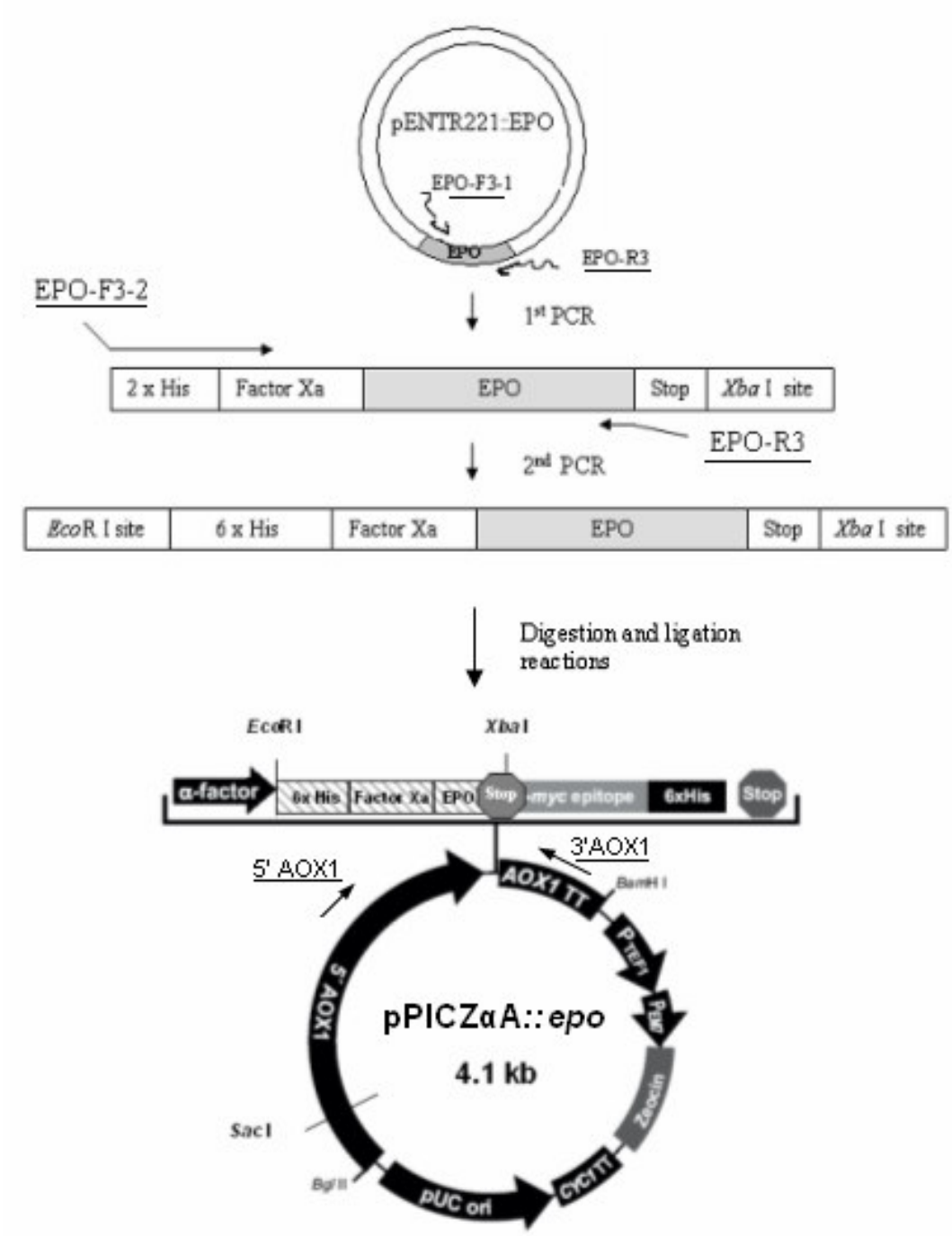


Figure 4.4 Schematic representation of epo amplification, integration of specific recognition sites by two-step-PCR, and construction of pPICZαA::*epo* plasmid. The *Eco*RI and *Xba*I sites used in ligation of the insert to the vector, the SacI site used to linearise the plasmid before transformation and the primers used for sequencing are also shown**.**

These criteria were also the reason for selection of pPICZαA vector rather than pPICZαB or pPICZαC, which only differ by their multiple cloning sites.

Six set of digestion reactions were performed during cloning steps:

1. pPICZαA digestion with *Xba*I for control
2. pPICZαA digestion with *Eco*RI for control
3. pPICZαA double digestion with *Xba*I and *Eco*RI for ligation
4. Insert DNA digestion with *Xba*I for ligation control
5. Insert DNA digestion with *Eco*RI for ligation control
6. Insert DNA double digestion with *Xba*I and *Eco*RI for ligation

As seen in lanes 1 - 4 of Figure 4.5, digestion of the vector DNA was successful at both ends. The single digestion at either end of the insert can not be visualized directly by agarose gel electrophoresis (Figure 4.5, lanes 5-6) since the fragment of DNA digested off by the enzyme is only a couple of nucleotides. Therefore, single site digestions of the insert DNA were ligated to itself as control of successful digestion and the presence of double sized insert DNA in ligation product verified its success (Figure 4.5, lanes 8-9). Other than the control ligations for digestion reactions, another control ligation reaction was set up, where there was no insert DNA. The main ligation reaction was performed to ligate the double digested insert DNA with double digested vector DNA as explained in Section 3.6.7. The product of this reaction is the putative plasmid named as pPICZαA::*epo* (Figure 4.4).

4.1.4 Transformation of *E. coli* cells with pPICZαA::*epo* and Selection of the True Transformant

To propagate the pPICZαA::*epo* plasmid, 5 µL of the ligation product was used in transformation of *E. coli* TOP10 cells as explained in Section 3.6.8.

Twelve single colonies randomly selected for further controls were grown in 5.5 mL LSLB medium containing 25 µg mL^{-1} Zeocin (Appendix E). 0.5 mL of 12 h culture was stored at 4°C, while from the rest; the plasmid was isolated (Section 3.6.1) and used in controls by restriction enzyme digestion and PCR. True positive colonies were plated (from the stored 0.5 mL culture) for short term storage, and further plasmid isolation to be used in control by DNA sequencing.

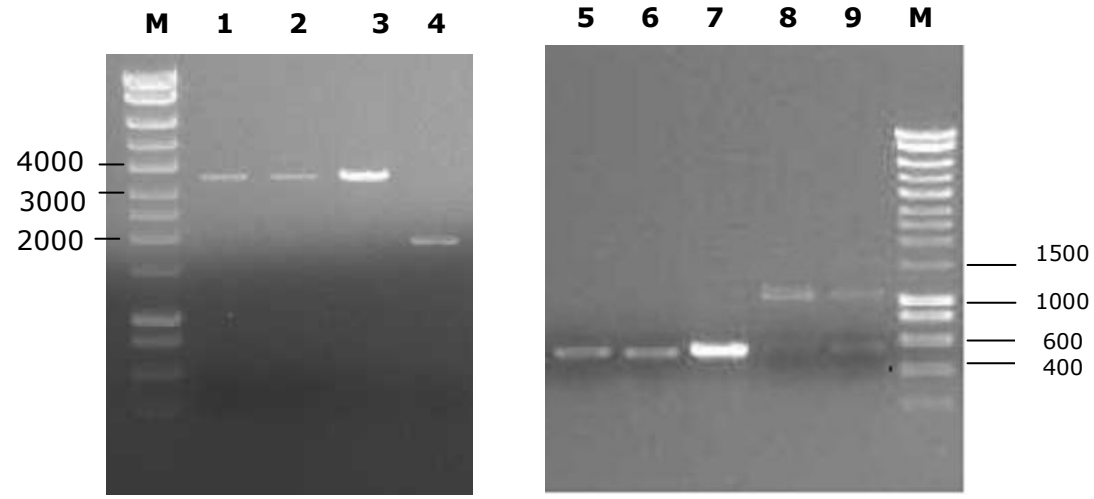


Figure 4.5 Agarose gel electrophoresis of products of digestion and ligation reactions. M: Hyperladder I, Lane 1: pPICZαA digested with *Xba*I, Lane 2: pPICZαA digested with *Eco*RI, Lane 3: pPICZαA digested with *Xba*I and *Eco*RI, Lane 4: Undigested pPICZαA, Lane 5: Insert DNA digested with *Xba*I, Lane 6: Insert DNA digested with *Eco*RI, Lane 7: Insert DNA digested with *Xba*I and *Eco*RI , Lane 8: Insert DNA digested with *Xba*I and ligated to itself, Lane 9: Insert DNA digested with *Eco*RI and ligated to itself.

The purified plasmids were first double digested with *Eco*RI and *Xba*I. As seen in Figure 4.6, six out of twelve colonies were shown to have an insert with the expected size when digested with same restriction enzymes used to clone the insert. The other six colonies were eliminated as false positives.

The selected six colonies were then controlled by PCR, using EPO-F3-2 and EPO-R3 primers, and all six colonies gave the 547 bp band as expected. Thereafter, plasmids from three of the selected six clones were selected for DNA sequencing (Section 3.6.9) using 5'AOX1 and 3'AOX1 primers. One of the three colonies gave the exact sequence (Appendix D) when analyzed by Basic Local Alignment Search Tool (BLAST). The sequencing primer set was sufficient to overlap and confirm the nucleotide sequence of the insert, as well as the α-factor signal sequence in the 5' end of the insert. The colony that gave the exact match of sequence was named as *E. coli* pPICZaA-EPO and its plasmid was named as pPICZαA::*epo* (Figure 4.4). Glycerol stocks were prepared for the selected clone.

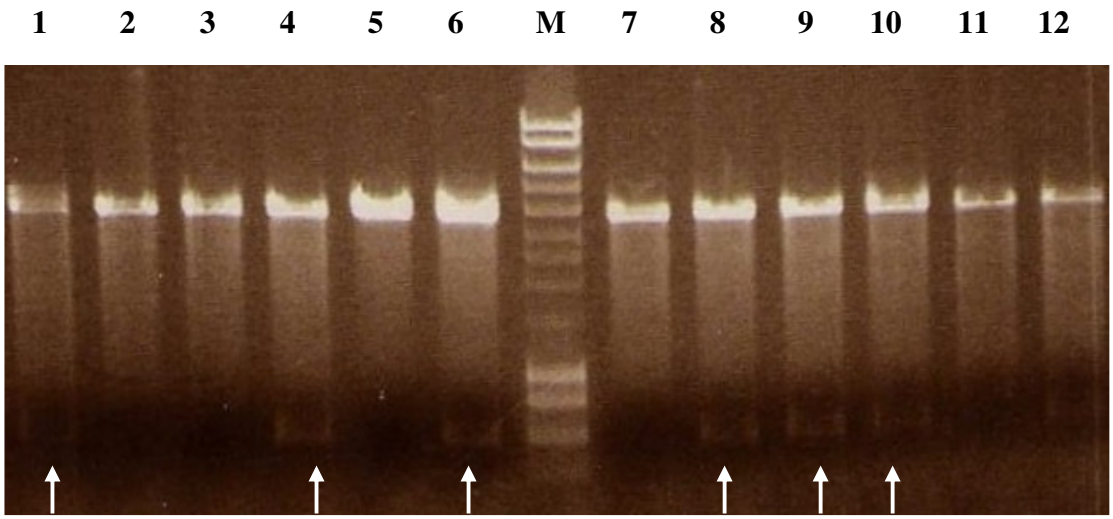


Figure 4.6 Agarose gel electrophoresis of double digestion by *Eco*RI and *Xba*I of the plasmids from the twelve selected *E. coli* transformants.

4.1.5 Transfection of *P. pastoris* cells with pPICZαA::*epo*

For the genomic integration of the pPICZαA::*epo* plasmid into the genome at the *AOX1* locus (Figure 4.7), the plasmid was linearized in its *AOX1* promoter region. For this purpose, *Sac*I was chosen as a single-cutter restriction enzyme in this region. Full digestion of the plasmid is important not to get false positives; therefore, digestion was carried on for 5 h and full digestion was verified by agarose gel electrophoresis. The digestion product was then purified (Section 3.6.6.1). The concentration of the digested plasmid solution had to be 0.1-0.2 μg μL^{-1} for efficient transfection and 50 μL was needed for each tube of transfection process. However, after purification, due to the high volume of water used for elution, the digested plasmid solution had to be concentrated by vacuum drying for 20-25 min. The concentration of the solution was then verified to be in the required range by Nanodrop™, which is an equipment requiring 1 μL sample and analysis is done spectrophotometrically. The digested, purified and concentrated plasmid DNA (pPICZαA::*epo*) to be used in transfection, was stored at -20°C till required.

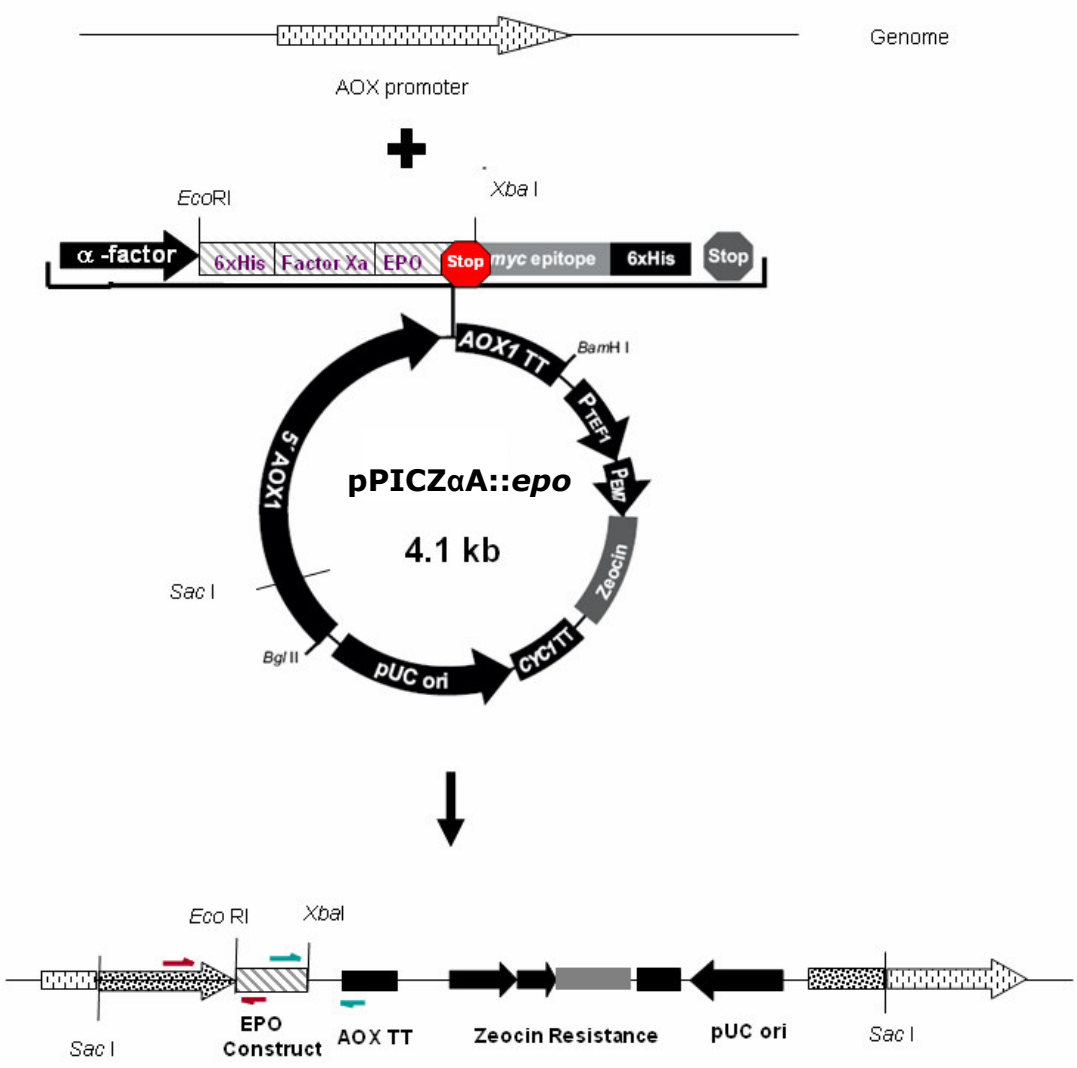


Figure 4.7 Schematic representation of pPICZαA::epo integration into *P. pastoris* genome. The plasmid, containing the EPO insert (box with dashed lines), was digested with *Sac*I from *AOX*1 promoter region, yielding a lianerized plasmid having homologous regions to the *AOX*1 promoter in the genome at both ends. After integration of the plasmid to the genome, there are two functional copies of the *AOX*1 promoter in the genome. The figure also shows the *Eco*RI and *Xba*I sites used in ligation of the insert to the vector.

Transfection was performed as explained in Section 3.6.10. Proper adjustment of concentration of each solution used in the protocol and being gentle to transfected cells in the resuspension step was observed to be important to obtain a high efficiency of transfection. Regardless of all the care taken, the highest transformation efficiency obtained was 160 cfu/µg linearized DNA, which was in lower end of the range proposed by the manufacturer (100 – 1000 cfu/µg linearized DNA).

After 60 h of incubation, twelve single colonies were selected from YPD + Zeocin (500µg mL^{-1}) plates for further controls. The selected colonies were restreaked onto YPD + Zeocin (1000µg mL^{-1}) plates for short term storage and then inoculated into 10 mL YPD medium containing 100 µg mL^{-1} Zeocin (Appendix E). Their genomic DNA was isolated to be used in controls by PCR and Southern blotting.

4.1.6 PCR and Southern Blotting Controls to Select the True Transformants

First control of integration of EPO gene to the *P. pastoris* genome, was done by PCR amplification, using 5'AOX1 and Seq-R3 as one primer set, expecting a product of 894 bp on agarose gel and 3'AOX1 and Seq-F3-2 as the second primer set, expecting a product of 708 bp (Figure 4.8).

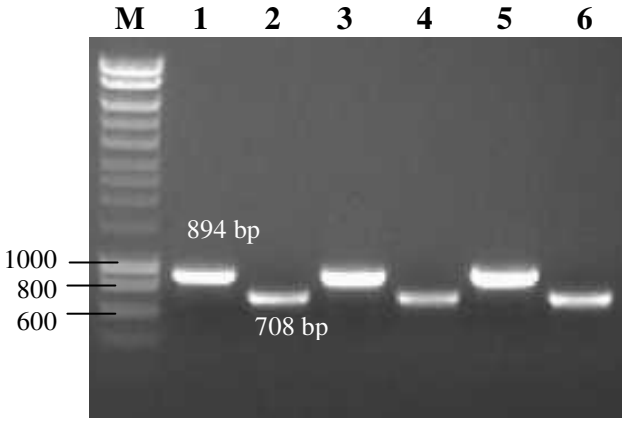


Figure 4.8 Agarose gel electrophoresis of PCR amplification of EPO insert from genomic DNA of *P. pastoris* transformants. Lanes 1,3, 5: 5'AOX1 and Seq-R3 primer set ; Lanes 2, 4, 6: 3'AOX1 and Seq-F3-2 primer set.

In Figure 4.8, the result of only three colonies is shown, though the same bands were obtained for all of the selected twelve colonies. The size of the bands being as expected, is an important proof towards genomic integration of the EPO gene. This procedure does not necessarily prove the integration of the gene to the desired locus in the genome; however, considering the large size of the homologous regions on the plasmid to the promoter of *AOX* gene, there is not much possibility of integration to another locus on the genome.

Due to unavailability of the genomic sequence of *P. pastoris* during this study, a primer complementary to non-plasmid DNA could not be designed. However, since the plasmid does not have an ARS (autonomously replicating sequence), the plasmid can not survive through generations, therefore could not give false positives.

A more specific proof of integration of the EPO gene to the genome was achieved by Southern blotting (Section 3.6.12) using the alkaline phosphatase labeled *epo* insert (512 bp) as the probe. Since with the Southern blotting probe, homologous recombination is much bigger in size and more specific than a PCR primer, the result of this analysis was a complete evidence of genomic integration of the EPO gene.

In Southern blotting, the genomic DNA digested with a single enzyme is run on agarose gel electrophoresis. *Bam*HI digests the pPICZαA::*epo* plasmid once outside the *epo* insert region and was a readily available enzyme. Wild type genomic DNA alone and together with plasmid DNA were run as controls, together with the twelve samples on agarose gel electrophoresis (Figure 4.9). Empty wells surrounded the samples so that the marker or the control plasmid would not leak and mix with the samples, giving false signals. The smear seen in Figure 4.9 is the expected view from the digested genomic DNA.

The band corresponding to *EPO* insert was seen in five colonies, as well as in the last control lane (Figure 4.10). The control lane, composed of digested plasmid DNA together with digested genomic DNA, gave a very light band, so the concentration of the plasmid DNA could have been increased for a better picture. The negative control (digested genomic DNA of wild type *P. pastoris*) did not give a band as expected (Lane 14, Figure 4.10). From the five colonies that gave the band in Southern blotting, glycerol stocks were prepared.

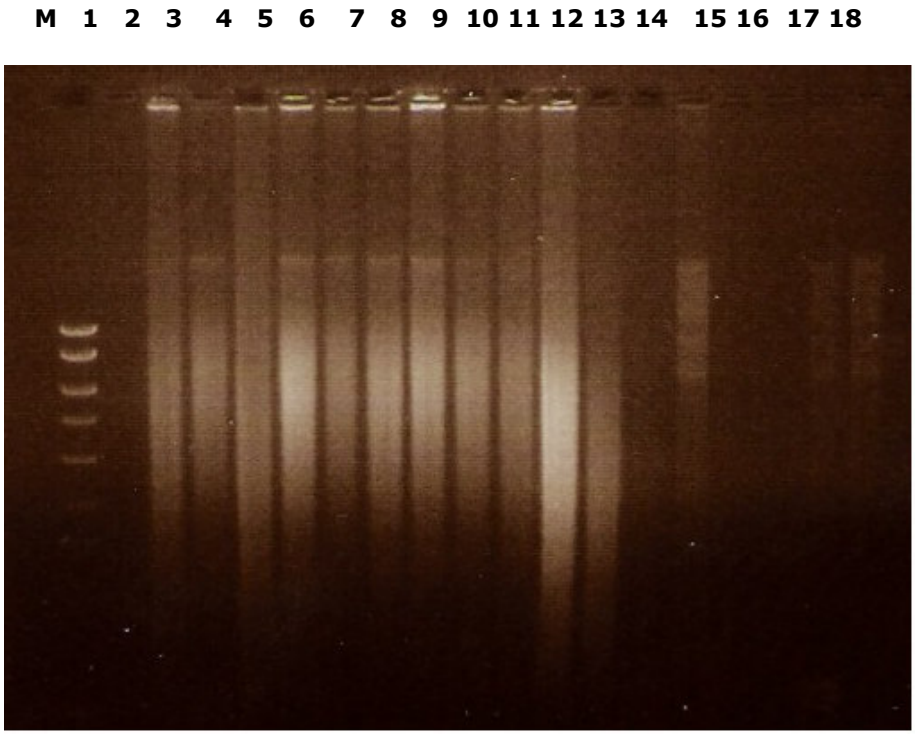


Figure 4.9 Agarose gel electrophoresis view of the *Bam*HI digested genomic DNA of *P. pastoris* transformants, to be used in Southern blotting. M: Hyperladder I, Lanes 1, 15,16: Empty wells, Lanes 2-13: *P. pastoris* transformants, Lane 14: Wild type *P. pastoris*, Lane 17: Wild type *P. pastoris* genomic DNA + 10-7 µg plasmid DNA, Lane 18: Wild type *P. pastoris* genomic DNA + 10-5 µg plasmid DNA.

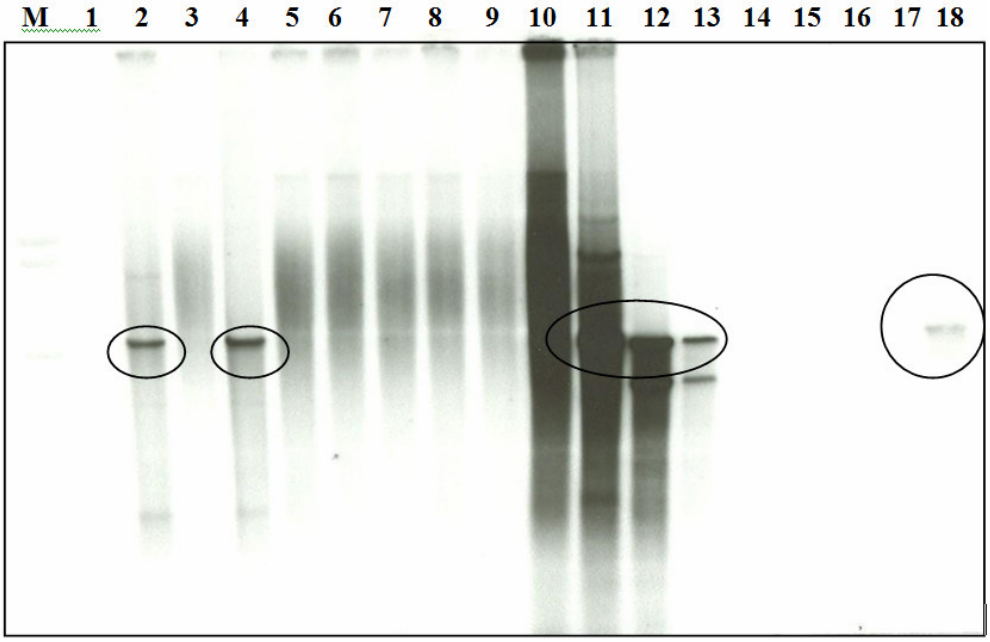


Figure 4.10 Southern blot hybridization of digested *P. pastoris* transformants' genomic DNA, using epo insert as the probe.

4.2 Expression and Purification of Recombinant Human Erythropoietin in *Pichia pastoris* for Structural Analysis

After selecting the true transformants by restriction digestion, PCR amplification and Southern blot analysis, the strain having the highest rHuEPO production capacity still had to be selected and this was performed by Western blot and mass spectrometry analyses. Thus, in this research phase, the potential strains were grown in a common complex medium, for the purpose of selection of the recombinant *P. pastoris* strain with the highest rHuEPO production capacity and thereafter for the structural analysis of the rHuEPO produced by this strain.

4.2.1 Western Blot Analysis for the Selection of Recombinant *P. pastoris* Strains with the Highest rHuEPO Production Capacity

The five selected colonies and the wild type *P. pastoris* as control were inoculated from glycerol stocks for protein production in 50 mL BMMY production medium, as described in Section 3.7.1.2. 20 mL of the samples were used in affinity purification (Appendix J), where 1 mL of resin was added directly into 20 mL supernatant and 300 µL of elution buffer was used to elute the recombinant protein. 20 µL of purified and 20µL of crude samples from each colony were run in SDS-PAGE and then the proteins on the gel were transferred to PVDF membrane for Western blotting analysis (Section 3.8.5).

The bands on Western blots appear through interaction of the antibody raised for EPO with the rHuEPO on the blotted Western membrane. Since the unglycosylated HuEPO is 18 kDa, the glycosylated rHuEPO is most likely the band with 30 kDa size. The smear on the gel is due to glycosylation. Although most of the protocols were not optimized for this study in the first run, the result clearly showed that two of the colonies had higher rHuEPO production capacities than the other three. The negative control was also successful as no band was obtained for the wild type (Figure 4.11).

The bands on Western blots appear through interaction of the antibody raised for EPO with the rHuEPO on the blotted Western membrane. Since the unglycosylated HuEPO is 18 kDa, the glycosylated rHuEPO is most likely the band with 30 kDa size. The smear on the gel is due to glycosylation. Although most of the protocols were not optimized for this study in the first run, the result clearly showed that two of the colonies had higher rHuEPO production capacities than the other three. The negative control was also successful as no band was obtained for the wild type (Figure 4.11).

The two transformants, *P. pastoris* E16 (Lane 3, Figure 4.11) and *P. pastoris* E17 (Lane 4, Figure 4.11), that gave similarly strong bands in Western blotting, also had growth profiles very similar to the wild-type *P. pastoris* in the production medium, therefore could not be distinguished (Figure 4.12).

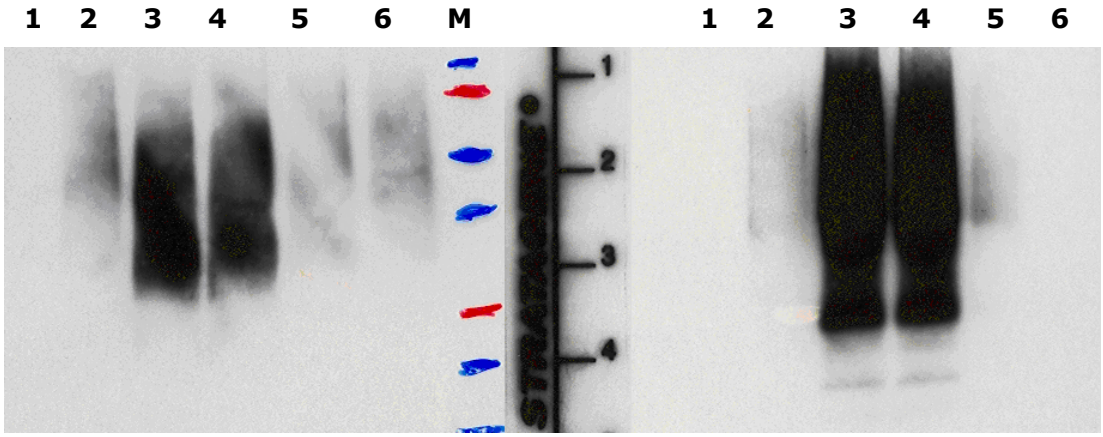


Figure 4.11 Western blot analysis of the proteins produced by *P. pastoris* transformants. M: Protein marker, marked from the gel onto the blot. Lane 1: Wild type *P. pastoris* as control. Lanes 2-6: The five colonies selected by Southern blotting. Left: Supernatant samples Right: Concentrated samples

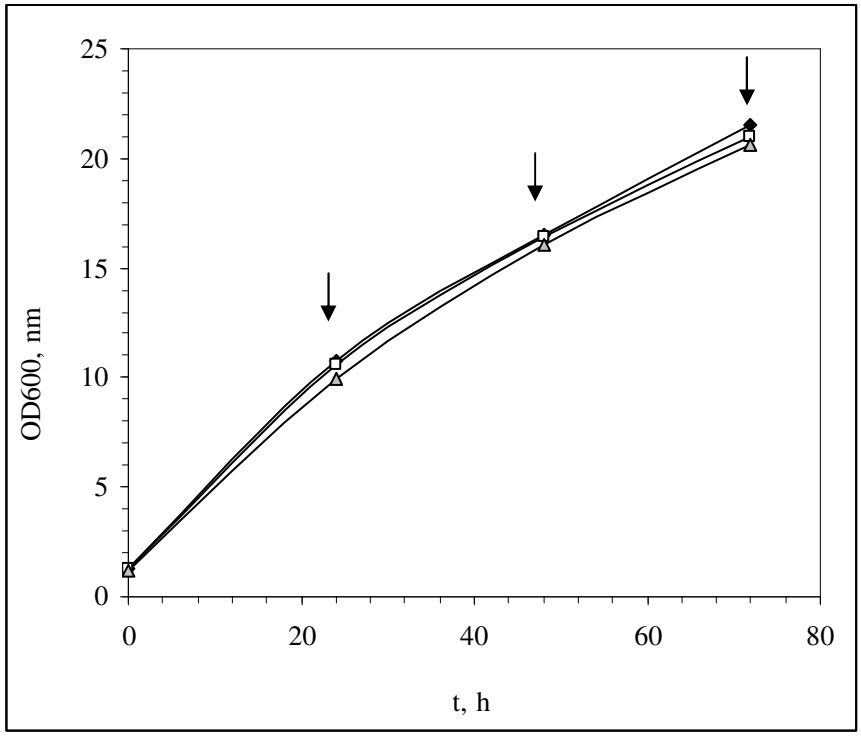


Figure 4.12 Optical density at 600 nm of the BMMY medium during production phase of *P. pastoris* cultivation. (♦) : *P. pastoris* wild type strain (▲) : *P. pastoris* E16 strain (□): *P. pastoris* E17 strain. Arrow indicates the time of methanol addition.

4.2.2 Q-ToF Mass Spectrometry Analysis for Confirmation of Primary Structure of rHuEPO

Recombinant protein expressed and purified from the two selected transformants, *P. pastoris* E16 and *P. pastoris* E17, were further analyzed. The 30 kDa band seen in SDS-PAGE gel (Figure 4.13, circled band) and Western blot, which was assumed to be the rHuEPO, was digested with trypsin and the peptides analyzed using a Q-ToF mass spectrometer (Section 3.8.6) with an electrospray ionization source to generate a peptide mass fingerprint. The predicted fragment products of the trypsin digestion reaction of rHuEPO are shown in Figure 4.14. The peptide fragments underlined in the figure were those identified by Q-ToF MS, based on a comparison of the calculated monoisotopic masses of the fragments to the measured masses obtained from the mass spectrum, typically accurate to 7 to 8 significant digits. The peptides identified from *P. pastoris* E17 had sequence coverage of 26% (that is, 26% of the protein

sequence was matched to the peptides detected in the MS), whereas those from *P. pastoris* E16 were likely not concentrated enough to be identified. More importantly, the peptide fragments detected include peptides close to both the *N*- and *C*- termini of the protein, indicating that expression is of the complete *epo* open reading-frame. The reason for the other fragments not being detected could be that either the fragments were too small or too large, or that they were glycosylated. The latter is especially expected in the fifth and ninth peptide fragments, where the glycosylation sites are present in HuEPO and possibly in rHuEPO.

The results of PCR analysis, Southern blotting, Western blotting and mass spectrometry analysis showed that *P. pastoris* E17 strain had the pPICZαA::*epo* integrated to its genome and that it was producing the rHuEPO. However, at this point, the 18 kDa native protein had still not been shown, which was the next step of the research.

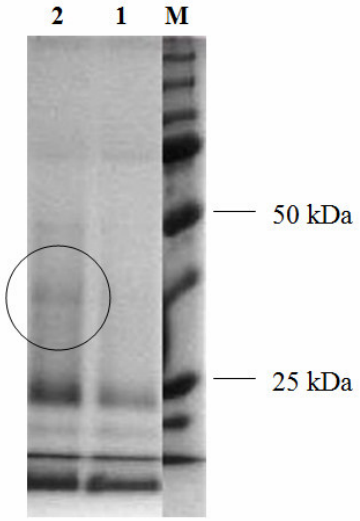


Figure 4.13 SDS-PAGE gel of *P. pastoris* production medium. The circled band was analyzed by mass spectrometry. M: Protein Marker (Appendix F), Lane 1: Proteins from *P. pastoris* E16 strain, Lane 2: Proteins from *P. pastoris* E17 strain.

appr licdsr vler ylleak eaen*ittgcaehcslnen*itvpdtk vnfyawk r

mevgqqavevwqglallseavlr gqallvn*ssqpweplqlhvdk avsglr slttllr

algaqeeaisppdaasaaplr titadtfr k lfr vysnflr gk lk lytgeacr tgd

Figure 4.14 EPO amino acid sequence. The bold letters (r and k) in the 165 amino-acid EPO sequence indicate the expected trypsin digestion sites. (Trypsin specifically hydrolyzes the carboxyl terminal side of arginine and lysine residues which do not precede proline residues.) The underlined peptide fragments are the ones detected in the Q-ToF mass spectrometry analysis. The star (*) after asparagine (n) residue indicates the glycosylation sites in EPO. The amino acid one-letter codes are given in Appendix G.

4.2.3 Purification and Obtaining the Native rHuEPO

As schematically explained in Figure 4.4, the recombinant protein produced by the *P. pastoris* expression system is not the native rHuEPO that is desired, but contains a polyhistidine tag for purification of the protein and a factor Xa recognition site at the *N*-terminal end of EPO. Thus, the *N*-terminal end can be digested from factor Xa recognition site following ultrafiltration and His-tag purification, therefore the native rHuEPO can be obtained.

4.2.3.1 Ultrafiltration

The first step of the purification process after cell growth and expression of the secreted recombinant protein was the concentration and desalting of the supernatant from the *P. pastoris* production medium, which was achieved by ultrafiltration as explained in Section 3.7.2. In choosing ultrafiltration membranes, as a rule of thumb, half the size of the protein should be selected as the molecular weight cut-off point. Since the unglycosylated form of the protein is 18 kDa, 10 kDa ultrafiltration membranes were used. Regenerated cellulose was preferred rather than polyethersulfone membranes, because they give the highest possible retention with the lowest possible adsorption of the protein.

During ultrafiltration and the following processes, it was important to keep the protein containing solution at 2-8°C to prevent microbial and proteolytic degradation. Therefore, the cold room was preferred during the process and the samples were kept on ice or in refrigerator. Dilute protein solutions of less than 1

mg mL^{-1} are more prone to inactivation and loss as a result of low-level binding to the storage vessel. Moreover, as the protein is concentrated, low amount of resin will be more efficient in His-tag purification. Therefore, ultrafiltration was carried till about 10-fold concentration is achieved. The next step was the polyhistidine–tag affinity purification process.

4.2.3.2 Polyhistidine–tag Affinity Purification

Cobalt based Immobilized Metal Affinity Chromatography (IMAC) resins (BD TALON™) were used for affinity purification of rHuEPO from the concentrated supernatant. The purification process had to be optimized though, for increased efficiency of separation. For instance, the pH of the protein solution directly affects the binding efficiency to the resin therefore has to be optimized. Moreover, the amount of resin, the incubation time, incubation temperature, and initial purity of the protein in the supernatant affect the efficiency of the purification process. Therefore, these parameters were optimized for higher efficiency.

First, the amount of buffer used in the washing step after ultrafiltration of the supernatant was increased to twice its volume, since this will increase the purity of the protein loaded into the affinity purification column. This increases the period of ultrafiltration and therefore increases the time for degradation of rHuEPO by proteases, even though the process is carried out at 4℃. Thus, the washing step after ultrafiltration should not be too long. Moreover, the wash buffer of affinity purification was used instead of pure water.

The effect of pH of the solution was investigated next. The isoelectric point (pI) of HuEPO is in the range, 3.92-4.42, and that of commercial rHuEPO has been reported to be, in the range 4.42–5.11 (Lasne and Ceaurriz, 2000). The reason for pI to be given in a range is due to microheterogeneity in the structure of glycoproteins. Isoelectric point is the pH where the net charge on the protein is zero. Therefore proteins are least soluble at their isoelectric point and they precipitate. The recommended pH for affinity purification is pH=7.0, thus wash buffers at pH 5.0, 7.0 and 8.0 were tried. Buffers with higher pH were not tried because at higher pH values, amino acids other than histidine would contribute to resin binding, therefore decrease selectivity. Five-fold decrease in efficiency compared to pH = 7.0 was obtained with buffer at pH=5.0. This is expected

since pH is closer to pI value, leaving less negative charge on the histidine side chain. Buffer with pH 8.0 showed no difference compared to pH 7.0, was therefore not preferred.

Thereafter step elution process was applied in order to get rid of other proteins with high histidine content, that have a low binding capacity though contaminate the actual protein to be eluted. Therefore, a pre-elution step was included, where the imidazole concentration in the elution buffer was 10 mM instead of 150 mM. The result of the step elution process can be seen in SDS-PAGE result in Figure 4.15, where proteins with much higher molecular weight (~70 kDa), therefore higher chance of histidine localization, seem to have been adsorbing to the resin together with rHuEPO. Thus, the pre-elution of unspecifically bound proteins contaminating the recombinant protein, was found to be important.

Figure 4.15 SDS-PAGE gel view of the proteins bound to the IMAC resins but eluted by step-elution prior to rHuEPO elution. Lane 1: Proteins eluted in pre-elution step M: Marker.

Finally the amount of resin used was increased, until all recombinant EPO was purified from the supernatant. Thus, two milliliters of resin was used for 20 mL ultrafiltered medium, which was obtained from 500 mL of production medium supernatant and the rHuEPO obtained was >5 mg per liter of production

medium, as calculated by the Bradford assay. This meant using four times more resin than the capacity stated by the manufacturer, due to the high amount of impurities still present in the supernatant.

The improvement in the purification efficiency can be seen by comparing SDS-PAGE gels before (A Lanes) and after (B Lanes) optimization studies (Figure 4.16). When Lanes A2 and B2 are compared, which were loaded with His-tag purified supernatant, before and after optimization studies, respectively, it can be seen that the rHuEPO band in Lane B2 has very low intensity. This is desired, since the rHuEPO should be adsorbed onto the resin, and should not be present in the supernatant, if the purification was efficient. Lane A3 shows the wash buffer passed over the resins during purification, which also should not contain the rHuEPO band, and it can be seen that this step is reasonably successful and not much recombinant protein is lost.

The rHuEPO obtained after optimization of the affinity purification process was >5 mg per liter of production medium, which showed about 5-fold improvement compared to purification prior to optimization studies. Therefore the optimizations for His-tag purification process were essential and successful. The 5 mg L^{-1} product obtained at this stage should be considered normal, since the aim at this stage was not mass production, thus the experiments were performed in laboratory scale shake bioreactors, without medium optimization and relatively low oxygen transfer conditions compared to large scale bioreactors.

4.2.3.3 Deglycosylation and Factor Xa Protease Digestion to Obtain the Native Protein

The purified rHuEPO was deglycosylated as explained in Section 3.7.5. To obtain a concentration of 5 µg of glycoprotein μL^{-1} of reaction buffer for the N-glycanase digestion, a buffer exchange was required and ultrafiltration spin columns (Sartorius AG, Göttingen, Germany) with a 10 kDa cut-off were used. The deglycosylated rHuEPO can be seen in Figure 4.17-a (lane 3) and Figure 4.17-b (lane 2), as having an apparent molecular weight of *ca.* 20 kDa.

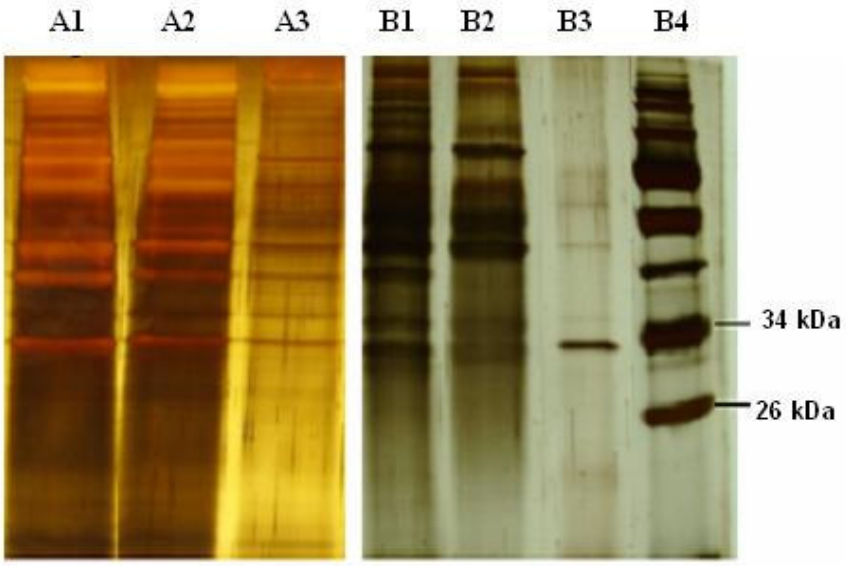


Figure 4.16 SDS-PAGE gel view before and after optimization studies. A1: Ultrafiltered production medium supernatant; A2: His-tag purified supernatant before optimization; A3: First wash buffer passed over the resins B1: Ultrafiltered production medium supernatant B2: His-tag purified supernatant after optimization B3: Double-purified rHuEPO B4: Protein ladder.

Finally, the polyhistidine tag was removed by factor Xa digestion (Section 3.7.4). The shift in the size of intact rHuEPO when digested with factor Xa protease can be seen from Lanes 1 and 2 in Figure 4.17-a. Factor Xa protease digestion site is after its recognition site, so there will not be any extra amino acids left following digestion. This was the reason for this protease to be chosen rather than other commonly used proteases such as thrombin, TEV protease and PreScission™ protease. The reason for preferring factor Xa protease instead of enterokinase, which also has a cleavage site at the end of its recognition sequence, was that the recognition sequence was shorter for factor Xa protease, and its recognition sequence was not present in EPO sequence. Moreover, EPO sequence did not have the (Gly-Arg) sequence, which is a secondary site of possible digestion; and did not begin with Pro or Arg, which prevents cleavage by factor Xa protease.

After removal of the 6xHis tag and the factor Xa recognition sequence by the factor Xa protease, the digestion mixture is passed through a special resin that efficiently removes the protease. This eluate is then passed through the metal

affinity resin to remove proteins not digested by factor Xa protease and the peptides containing the 6xHis tag.

The finally purified, deglycosylated and factor Xa-digested rHuEPO, corresponding to the wild-type polypeptide, when analyzed by SDS-PAGE (Figure 4.17-a, lane 4) and Western blotting (Figure 4.17-b, lane 3), showed an apparent molecular mass of the expected size 18 kDa. The constructed expression plasmid thus provides the potential of large-scale production of recombinant human EPO and its processing to give the native form. This is more advantageous compared to the studies, where rat EPO cDNA cloned directly into pPICZαA for expression in recombinant *P. pastoris* (Hamilton et al. 2006), and where rEPO expressed in *Drosophila* S2 cells, could not be obtained in native form following affinity purification, since the fusion construct including the polyhistidine tag was not digestable. The relatively lighter bands seen in silver stained gel electrophoresis view are the incompletely digested proteins, which can be removed with prolonged digestions, and unspecifically bound proteins in polyhistidine-tag affinity purification, which can be removed by step elution.

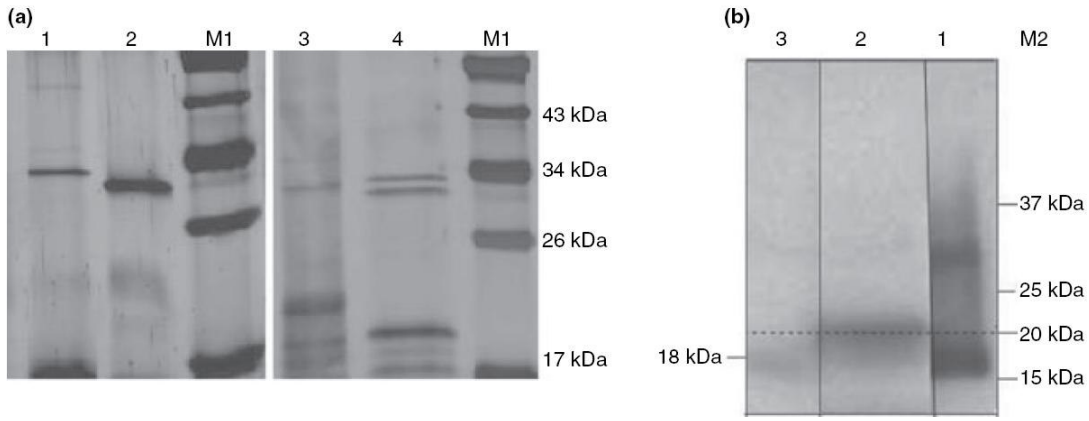


Figure 4.17 (a) Silver stained SDS-PAGE view of rHuEPO produced by *P. pastoris* E17. Lane1: Polyhistidine-tag affinity purified rHuEPO from *P. pastoris* E17 production medium. Lane 2: Purified and factor Xa digested rHuEPO. Lane 3: Purified and deglycosylated rHuEPO. Lane 4: Purified, deglycosylated, and factor Xa digested rHuEPO. M1: PageRuler Prestained Protein Ladder (Fermentas, Germany). **(b)** Western blot view of rHuEPO produced by *P. pastoris* E17. Lane 1: Polyhistidine-tag affinity purified rHuEPO from *P. pastoris* E17 production medium, Lane 2: Purified and deglycosylated rHuEPO, Lane 3: Purified, deglycosylated, and factor Xa digested rHuEPO. M2: Positions of protein size marker (Presicion Plus dual color prestained protein marker, BioRad, Hercules, CA).

4.2.4 Structural Analysis by MALDI-ToF MS

EPO is a naturally heavily glycosylated (40% of total mass) protein, and the glycan structure is important for its structure, function and stability (biological half-life) (Higuchi et al., 1992). RHuEPO produced by *P. pastoris* E17 in this study is also glycosylated by the post-translational biosynthetic processes of the yeast host, but these glycosylating modifications differ from those in human cells. The difficulty in glycoprotein analysis is that a large number of techniques are required for a full characterization of the structure (Anumula, 2000) and there are many ways to characterize the glycosylation of glycoproteins, including their structure. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-ToF MS) and electrospray ionization mass spectrometry (ESI-MS) are the most sensitive and often-used methods of mass measurement analysis due to their "soft ionization" techniques (Morelle and Michalski, 2005). MALDI-ToF MS is the preeminent technique for glycan analysis because it is more sensitive compared to ESI-MS (Dell and Morris, 2001), unlike many other forms of analysis, spectra can be obtained from unmodified glycans (Harvey, 1999), and only picomole amounts of oligosaccharide sample are needed (Küster et al., 1997). Therefore, the information involving the amount of *N*-linked glycans and the level of mannosylation of rHuEPO were simply determined by analyzing the products of the deglycosylation reaction using MALDI-ToF MS.

4.2.4.1 Initial Optimization Studies

Initial optimization studies had to be performed, as there are many factors affecting the signal/noise ratio. These are mainly:

- Purity of the sample
- Matrix/sample ratio
- Spotting technique
- Voltages
- Laser firing energies and duration on a spot

Purity of the sample was observed to be very important, however many impurities comes from the previous steps. Therefore, the sample was passed through spin column and washed with double distilled water many times or drop dialyzed for at least 40 min and then dried with centrifugal evaporator. The main problem is the SDS in the samples. In Figure 4.18, the difference between a well crystallized sample spot (Figure 4.18-B) and one with excess SDS (Figure 4.18-C) can be seen. The presence of SDS can be detected when the laser fires the spots, the sample shines like a mirror. Interestingly, the amount of SDS used prior to deglycosylation was found to be crucial in determining the dominant form of glycan. This is probably due to suppression of ionization by SDS, which can not be easily removed after the deglycosylation reaction.

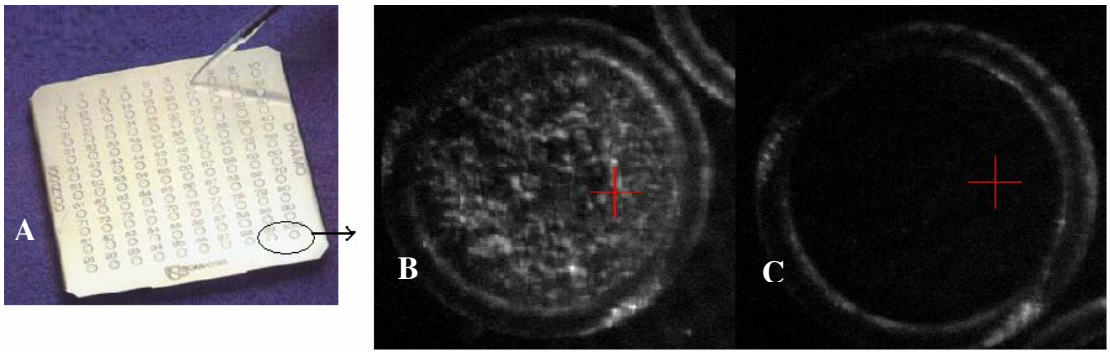


Figure 4.18 A Target plate and magnified photographs of the sample spots for MALDI-ToF MS. A. Target plate with 100 spots, where size can be compared to the size of a pipet tip B: Well crystallized sample spot C: Uncrystallized sample spot due to excess SDS.

Matrix and sample concentrations and matrix/sample ratio should be within detectable limits of MALDI-ToF MS which is theoretically in the femtomole levels. Nevertheless picomole amounts of the sample were generally needed to be applied, to obtain a good signal/noise ratio. Matrix:sample in 3:1 ratio was optimal in almost all trials. Matrix has to be abundant for good ionization of the sample, since it absorbs the laser energy and transfers it to the sample.

Amount of acetonitrile in matrix solution was observed to affect the crystallization properties. Higher acetonitrile concentrations (50%) were

preferable for protein analysis, whereas the suggested acetonitrile concentration in literature for glycan analysis was 30% and it was not changed.

There are many techniques in literature for spotting the sample and matrix on the target plate, which affects the crystallization and thus the energy transferred from the matrix to the sample. Spotting should be practiced carefully because samples with 1-2 µl volumes are spotted, which has to be fast or the mixture will evaporate, changing the concentration and thus crystal properties. The easiest technique was the so called "dried droplet" technique where the matrix and sample are mixed prior to spotting. But "layer by layer" technique where the sample is sandwiched between two layers of matrix was more efficient. Moreover, in the case of glycans, relayering the crystallized spots with 0.3-0.5 µl ethanol increased spot to spot reproducibility, as finer and more homogenous crystal layer was formed. Excess ethanol caused the sample to leak out of the well.

The voltages affect whether the detector can catch the ions in the order which they can be distinguished. They can be optimized per each protein or glycoprotein studied. However optimizations are usually rather done for the type of molecule studied, whether it is a protein or glycan.

The most time consuming optimization for any new molecule to be studied is probably the laser firing energy and learning the so-called "sweet spots" on the sample and then selecting the spectra which have been generated from only the sweet spots. Usually about a hundred of spectra have to be generated from a spot to be statistically correct, and then these are combined. Laser firing energy needed for glycans to be ionized was found to be much higher than that needed for proteins and the spots would burn out faster. Therefore, the laser has to be moved around faster and systematically for glycans spectra to be more precise when combined. Otherwise only 10 out of 100 spectra would serve useful and others would just diminish the peaks when spectra are combined.

4.2.4.2 The Main Results

Neutral glycans produce predominantly $[M+Na]^+$ ions during MALDI ionization (Harvey, 1999). The most intense isotopic peak cluster seen in the mass spectrum of the glycans of rHuEPO shown in Figure 4.19**-**A is *m/z* 3204, which

corresponds to $Man_{17}(GlcNAc)_2$, and is denoted as M_{17}. The isotopes in the cluster are separated by 1 unit on the *m/z* axis, *z* = 1 and *m* = 3204 Da, which is practically the same as the calculated $[M+Na]^+$ ion mass, 3203.6 Da. Glycosylation is known to be a posttranslational modification with considerable microheterogeneity, meaning that a glycoprotein is synthesized as a group of glycoforms (Papac et al., 1996). So, as expected, there are also peaks present in the mass spectrum of relatively lower mass, corresponding to other high mannose-type glycans, particularly M_9, M_{10} and $M_{11.}$

It has been shown that there is a linear relationship between the signal produced by the $[M+Na]^+$ ion from neutral glycans and the amount applied to the target, irrespective of structure (Harvey, 1993; Harvey, 2005), unlike the case for peptides and glycopeptides, where signal intensity depends on the proton affinity of the compound (Naven and Harvey, 1996; Harvey, 2005). Therefore, $Man_{17}(GlcNAc)_2$ (3.2 kDa) can be said to be the dominant form of *N*-linked glycosylation in rHuEPO, since the highest signal in Figure 5A is obtained from M_{17}. *P. pastoris* is known to produce oligosaccharides with 8-18 mannose residues as the major modification component (Hirose et al., 2002). Further, with the use of other wild type *P. pastoris* strain (NRRL-Y11430) producing rHuEPO, 9-13 mannose residues were detected (Hamilton et al., 2006).

The presence of SDS, used for the purpose of unfolding the glycoprotein substrate so that the N-Glycanase has greater access, had the unwanted effect of suppressing the ionization of the dominant glycan form in MALDI spectral analysis. When the final concentration of SDS in the deglycosylation reaction was 0.1%, which accorded to the manufacturer's recommendations and has been used in many other studies, the glycans $M_{9\text{-}11}$ were detected but M_{17} was never seen (Figure 4.19-B). However, when 10-fold less SDS was used in the deglycosylation reaction (0.01% SDS of final concentration), the most spectrally intense glycan detected was M_{17} (Figure 4.19-A). Although in both cases drop dialysis was carried out for 45 min to remove any low molecular weight substances including SDS, the results show that the higher mass glycan form was nonetheless undetectable by MALDI-ToF MS.

MALDI MS analysis of the N-glycanase-treated rHuEPO purified from the medium of the recombinant *P. pastoris* (not digested with factor Xa protease) shows a spectral peak centered about *m/z* 20400 (Figure 4.20). Production of singly

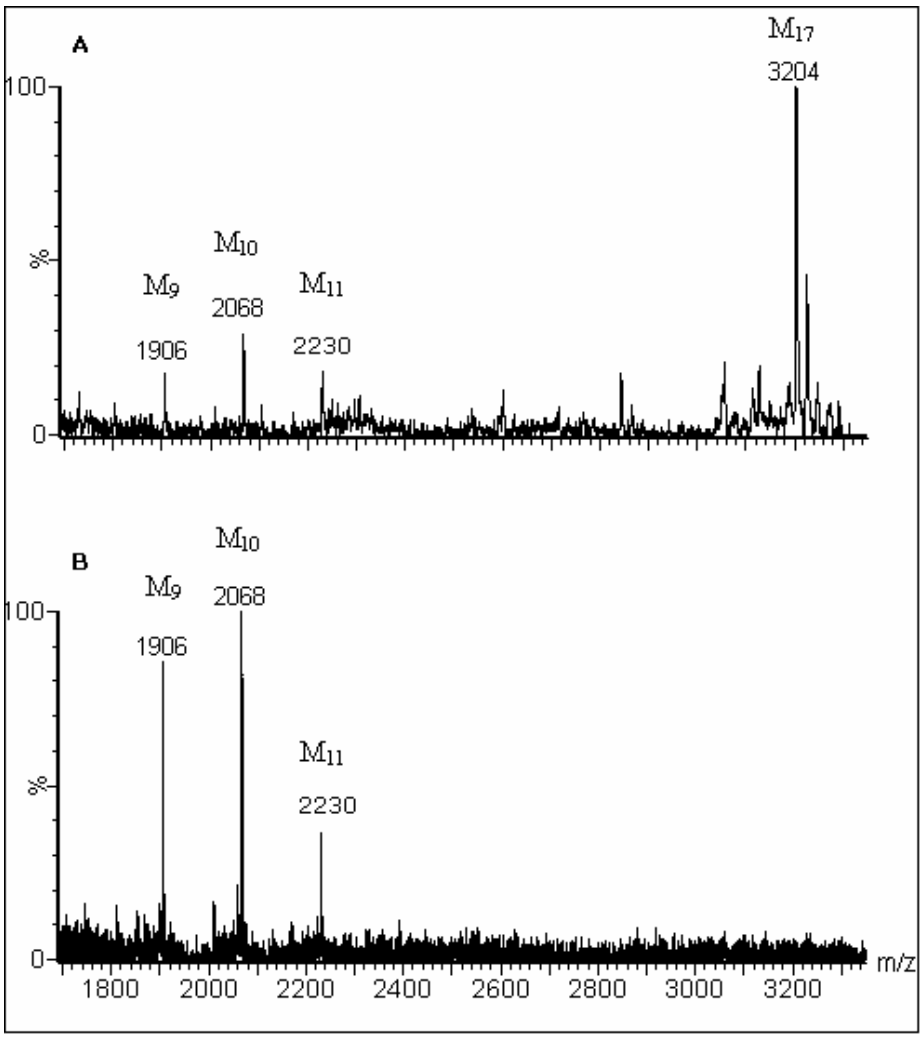


Figure 4.19 MALDI mass spectra of glycans of rHuEPO (released by N-glycanase treatment) on an S-DHB matrix, detected in $[M+Na]^+$ form. Final concentration of SDS used in protein denaturation before deglycosylation reaction: **A.** 0.01% **B.** 0.1%. M_9: $Man_9(GlcNAc)_2$; M_{10}: $Man_{10}(GlcNAc)_2$; M_{11} : $Man_{11}(GlcNAc)_2$; M_{17} : $Man_{17}(GlcNAc)_2$.

charged (z = 1) ions of intact proteins is typical in MALDI ionization, and so this *m/z* value agrees quite well with the 20 kDa apparent molecular weight of the band detected in a Western blot loaded with N-glycanase-treated rHuEPO (Figure 4.18-b, lane 2). A relatively low intensity peak in the spectrum, present at *m/z* 29700 (Figure 4.21), would correspond well with a small amount of intact rHuEPO glycoprotein that had not been deglycosylated with N-glycanase treatment (Figure 4.17, lane 1). Moreover, the second low-intensity peak in Figure 4.20 is most likely the partially deglycosylated rHuEPO, which has lost one of its *N*-linked glycans, as the molecular weight difference between intact protein of 29.7 kDa and that of 26.6 kDa peak corresponds to the molecular weight of a single *N*-linked glycan, perhaps M17. The spectral peak at *m/z* 35740 (Figure 4.20) likely corresponds to the mass of the N-glycanase (PNGase F) itself, which has a reported molecular weight of 35500 Da (Tarentino et al., 1985). Thus, the enzyme was detected with 0.7% error and therefore served as an internal mass calibrant. The molecular weight difference between the deglycosylated rHuEPO (*ca.* 20.5 kDa) observed in MALDI and of the native protein (*ca.* 18 kDa) is attributable to the amino-acid residues on the *N*-terminal portion of the polypeptide that had not been removed by factor Xa protease digestion, as shown in Figure 4.17.

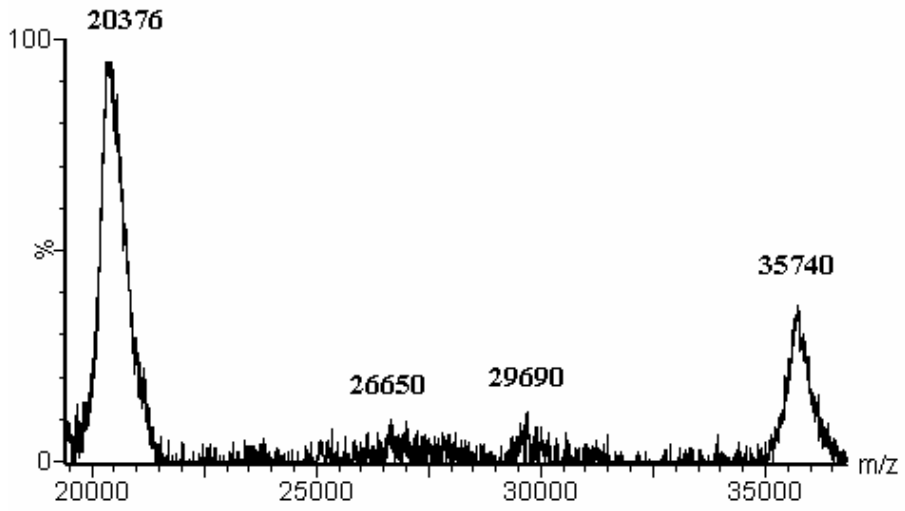


Figure 4.20 MALDI-ToF mass spectrum of the deglycosylated rHuEPO (m/z=20376) on sinapinic acid matrix. Due to incomplete reaction, the glycosylated rHuEPO (m/z=29690) and partially deglycosylated rHuEPO (m/z=26650) were also detected. The peak at m/z 35740 is the N-glycanase enzyme left from deglycosylation, which also serves as an internal mass calibrant. All the molecules detected are in singly charged, $[M+H]^+$ form.

The RHuEPO produced by *P. pastoris* E17 had a molecular weight *of ca.* 30 kDa, and the deglycosylated form was *ca.* 20kDa, which were detected both by Western blotting and MALDI-ToF MS. Since EPO has three possible sites for *N*-glycosylation, it is most probable that each *N*-glycan has a molecular weight of about 3.1-3.2 kDa, as derived from the molecular weight differences between the glycosylated- (*ca.* 30 kDa; Figure 4.17-B, lane 1 and Figure 4.20) and deglycosylated- (*ca.* 20.5 kDa; Figure 4.17-B, lane 2 and Figure 4.20) rHuEPO, thus also proving that $Man_{17}(GlcNAc)_2$ should be the dominant form of *N*-linked glycosylation in rHuEPO. Moreover, because the molecular weight difference between the glycosylated and deglycosylated rHuEPO is more than 9 kDa and when divided between three possible sites for *N*-glycosylation, $M_{9\text{-}11}$ (with molecular weights between 1.9 – 2.2 kDa) would not be expected. This result demonstrates the importance of considering the possibility that the absence of expected spectral peaks in MALDI MS might be due to effects of ion suppression that cannot be entirely overcome by standard sample preparation or clean-up. Thus, the already complicated glycosylation profiles of many glycoproteins could be determined incorrectly due to presence of ionization inhibitors like SDS or other salts.

Mass spectrum of the deglycosylated and factor Xa digested rHuEPO is shown in Figure 4.21. The major ionization form of the protein with the sinapinic acid matrix used, was the $[M+2H]^{2+}$ form, thus z = 2 with a peak at *m/z* position of 9.2 kDa, which means *m* is around 18.4 kDa. This result is in agreement with the findings of Stanley and Poljak (2003), where they analysed deglycosylated EPREX (commercial rHuEPO, expressed in mammalian cells), and also correlates well with the SDS-PAGE and Western data in Figure 4.17. Moreover, the larger peak at *m/z* position of 10 kDa is likely the $[M+2H]^{2+}$ form of rHuEPO which has been incompletely digested with factor Xa protease, thus when multiplied by the charge, *z* = 2, mass of 20 kDa corresponds well with the 20 kDa peak in Figure 4.20. Having detected the multiply charged molecules is advantageous, as these values are more accurate than those of the singly charged ions because of better statistics (Sttübiger et al., 2005), where the charge on the molecule is dependent on the choice of the matrix and its interaction with the analyte.

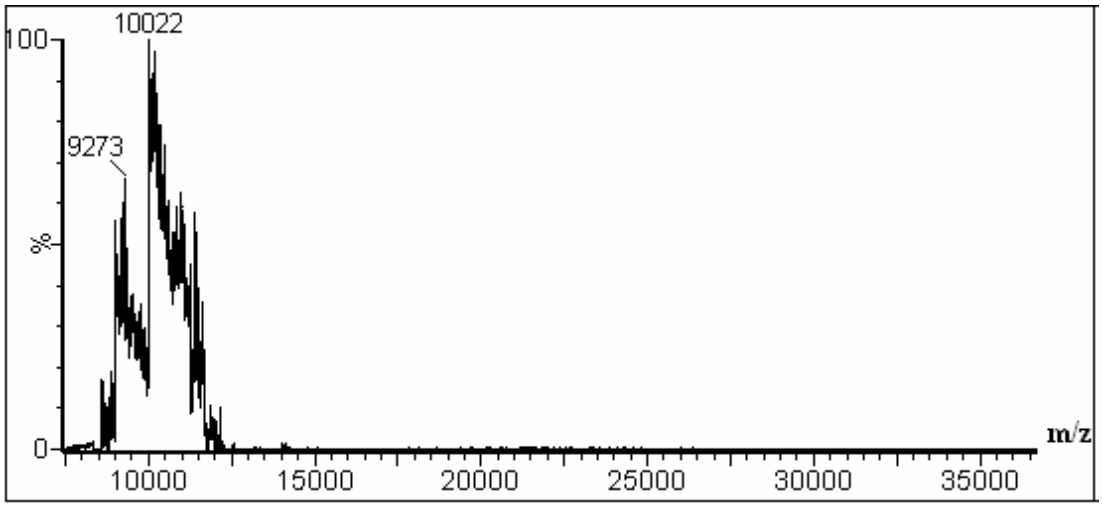


Figure 4.21 MALDI-ToF mass spectrum of the native rHuEPO, deglycosylated and factor Xa protease digested rHuEPO, (m/z=9273) on sinapinic acid matrix, detected in $[M+2H]^{2+}$ form. Due to incomplete digestion reaction with factor Xa protease, the $[M+2H]^{2+}$ form of only deglycosylated rHuEPO is also detected (*m/z* =10022).

The *O*-linked glycans of yeasts are relatively small (i.e., Man_{1-6} in *S. cerevisiae*); they cannot be digested by enzymatic methods and, moreover, the chemical methods employed for digestion may cause a 'peeling reaction' that destroys the oligosaccharide species (Gemmill and Trimble, 1999). Thus, little has been reported concerning the *O*-glycans of *P. pastoris*, and they were not analyzed in this study. Furthermore, their role in functioning and stability of EPO in human has not been defined and is, perhaps, not crucial.

4.3 Expression of Recombinant Human Erythropoietin in *P. pastoris* for Optimization Studies in Laboratory Scale Air Filtered Shake Bioreactors

After the recombinant microorganism producing rHuEPO was developed and the recombinant protein produced by the microorganism was purified and characterized by Western blotting and MALDI-ToF MS, the effects of medium components and pH on rHuEPO production and cell growth were investigated in laboratory scale air filtered shake bioreactors. Medium design studies were carried out in two distinct parts, where both complex and defined medium components were investigated.

4.3.1 Complex Medium Design

An alternative and cheap carbon source for *P. pastoris* fermentation processes was suggested for the first time in this study, that is using the byproduct of biodiesel industry, crude glycerol.

In spite of the high biodiesel production potential that triggers high glycerol liberation, there is no work in the literature on the use of crude glycerol in fermentations by methylotrophic yeast, eliminating the purification step of methanol, which is generally inhibitory for cell growth. Thus, in the present work, we have investigated for the first time, the effects of crude glycerol including methanol, obtained from the transesterification reaction of different vegetable oils namely, canola, sunflower, soybean, and corn oils, on recombinant protein production using recombinant *Pichia pastoris.*

4.3.1.1 Glycerol from Biodiesel via Transesterification of Various Vegetable Oils

Biodiesel was synthesized by transesterification reaction using canola, sunflower, soybean, and corn oils, at various oil:methanol:NaOH molar ratios. The oil:methanol:NaOH molar ratios for both soybean and corn oils were 1:6:1. The reaction conditions for the canola oil (C) and sunflower oil (S) used in the production medium are given in Table 4.1, which also shows the corresponding sample codes besides the glycerol and methanol contents of the glycerol phase used in the bioprocess. The biodiesel synthesis process can be represented simply as:

Glyceride (vegetable oil) + **n** Methanol $\rightarrow$ **n** Ester (Biodiesel) + Glycerol

Biodiesel yields were calculated from the statistical model (Gemma et al., 2007) and are presented in Table 4.1. The catalyst concentration, temperature and vegetable-oil:methanol molar ratio apparently have significant effects on biodiesel yield.

Table 4.1 Transesterification reaction conditions for biodiesel synthesis, using canola oil (C) and sunflower oil (S)[a].

Transesterification reaction Oil:Methanol:NaOH	Biodiesel Yield (%)	Glycerol in the sample (g L^{-1})	Methanol in the sample (g L^{-1})	Sample Code
1:6:1	94	456	478	S1
1:6:0.5	97	627	652	S2
1:3:0.5	83	881	544	S3
1:6:1	94	293	288	C1
1:6:0.5	97	478	499	C2
1:3:0.5	83	646	473	C3

a. The biodiesel yield, the glycerol and methanol contents of the glycerol phase (sample) used in the fermentation process as feed, and their corresponding codes are also given.

4.3.1.1.1 Preparation of the Glycerol Phase of the Transesterification Reaction for Fermentation Process

The glycerol samples, separated from the biodiesel layer by the separatory funnel, were dissolved in distilled water, such that the glycerol concentration of each stock solution was 126 g L^{-1}. The filter sterilized stock solutions were then mixed with the rest of the production medium components, so that the initial glycerol concentration was C_{G0}=12.6 g L^{-1} in all the media. At this point, the white-oily precipitates formed in the media, potentially the potassium and sodium salts of free fatty acids, was separated by centrifugation at 10000 *g*, 4℃ for 30 min. Clear solutions were obtained for the samples produced using canola and sunflower oil. On the other hand, even after 1 hour of centrifugation, the free fatty acid layer that contains fines of the salt did not precipitate out. As obtaining a clear solution is necessary for proper spectrophotometric measurement of cell concentration, the glycerol samples obtained from transesterification reaction using soybean and corn oil were eliminated at this stage, as being inefficient production medium components. Thus, canola and sunflower oils were used in further experiments (Table 4.1). Finally, the pH of the media were controlled and found to deviate from the pure glycerol medium by maximum 4%.

4.3.1.2 The Bioprocess by *P. pastoris* for Recombinant Protein Production using Crude Glycerol

P. pastoris was chosen as the recombinant protein producing host microorganism, using the byproduct of biodiesel production due to its toleration limit to high methanol concentrations (as the availability of the AOX1 enzyme promotes conversion of the toxic methanol into formaldehyde) and due to the fact that glycerol and methanol have been the widely accepted industrial choice as the carbon sources for *P. pastoris* fermentations. Moreover, methanol induces the recombinant protein expression under the control of the *AOX*1 promoter of *P. pastoris.*

After selection of the crude glycerol samples obtained from transesterification reaction of canola and sunflower oils, the effects of using glycerol originating from different oil sources as well as different preparation techniques of the biodiesel (the differences explicitly given in Table 4.1) on the bioprocess and on recombinant protein production characteristics were investigated.

4.3.1.2.1 Effect of Biodiesel Byproduct Glycerol on Substrate Consumption Profiles

Glycerol is regularly used as the main initial carbon source in *P. pastoris* fermentations to increase the cell concentration. During this first period (cell generation) of the bioprocess, cells prefer to consume solely glycerol and biomass accumulates, but recombinant protein production is fully repressed. Thereafter, methanol is fed to the culture to induce expression of the recombinant protein under the control of *AOX*1 promoter (Cereghino and Cregg, 2000).

The efficiency of using glycerol samples, supplied as the main byproduct of biodiesel production (instead of pure glycerol), in the production medium was investigated by comparing the substrate (glycerol) concentration profiles of *P. pastoris* E17 at t=0-30 h, which corresponds to the glycerol consumption period (Figure 4.22). Glycerol concentration decreased with the cultivation time, being totally consumed at t=30 h in all the conditions. The substrate consumption rates for different glycerol samples varied (with $p<0.05$), as calculated from Figure 1; and interestingly, with pure glycerol (sample G1), the lowest initial

glycerol consumption rate ($r_G = dC_G/dt$) of $-r_G$ = 3.2 mol m^{-3} h^{-1} (0.29 g L^{-1} h^{-1}) was obtained, whereas with sample C3 (Table 4.1) the highest value ($-r_G$ = 5.0 mol m^{-3} h^{-1}) was attained, being 1.6-fold higher than that obtained with pure glycerol. This is likely due to the presence of fatty acids, vitamins A, E and K (Heinonen et al., 1997; Gao et al., 1995) and trace elements (Cindric et al. 2007) in the vegetable oils, diffusing to the glycerol phase during the biodiesel formation reactions, and thus enriching the glycerol based production medium. Evidently, these compounds have various positive effects on the yeast physiology and biotechnology such as, improved membrane integrity and ethanol tolerance (Walker, 1998), increase in intracellular NAD level (Chen et al., 2007) and as antioxidants (Emri et al., 2004). Moreover, a detailed study on the macro-elemental analysis of various crude glycerol revealed the carbon, nitrogen, calcium, magnesium, phosphorus, sulfur and sodium contents, having average values of approximately 25% , 0.05%, 20 ppm, 5 ppm, 50 ppm, 15 ppm, 1%, respectively (Thompson and He, 2006), where the components have an undoubted positive effect on yeast metabolism.

4.3.1.2.2 Effect of Biodiesel ByProduct Glycerol on Cell Growth Profiles

The efficiency of using glycerol samples in the production medium was investigated by comparing the cell growth profiles of *P. pastoris* E17 for the t=0-54 h of the bioprocess (Figure 4.23). At t = 54 h, cell concentration (C_X) reached the maximum value, C_X = 10.5 g L^{-1}, in the medium containing glycerol sample C1; being significantly higher than the other carbon sources ($p<0.05$). On the other hand, in the medium with pure glycerol (G1), having the lowest glycerol consumption rate, the lowest cell concentration, C_X = 7.1 g L^{-1}, was obtained. However, as there exists other factors affecting the cell growth, the glycerol sample C3 with the highest glycerol consumption rate could only reach to C_X = 7.7 g L^{-1}, at t=54 h. As well as the additional nutrients coming from the vegetable oils, which possibly increases the cell growth rates as discussed previously, residual methanol in the glycerol phase also affects the metabolism.

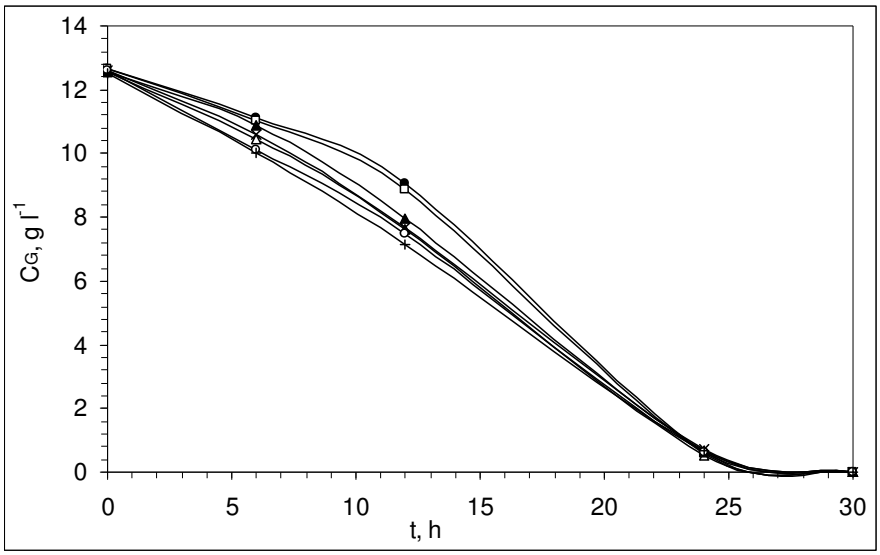


Figure 4.22 Variations in the glycerol concentration with the cultivation time in media containing different sources of glycerol: (●) pure glycerol (G1); (Δ) S1; (□) S2; (x) S3; (☐) C1; (O) C2; (+) C3.

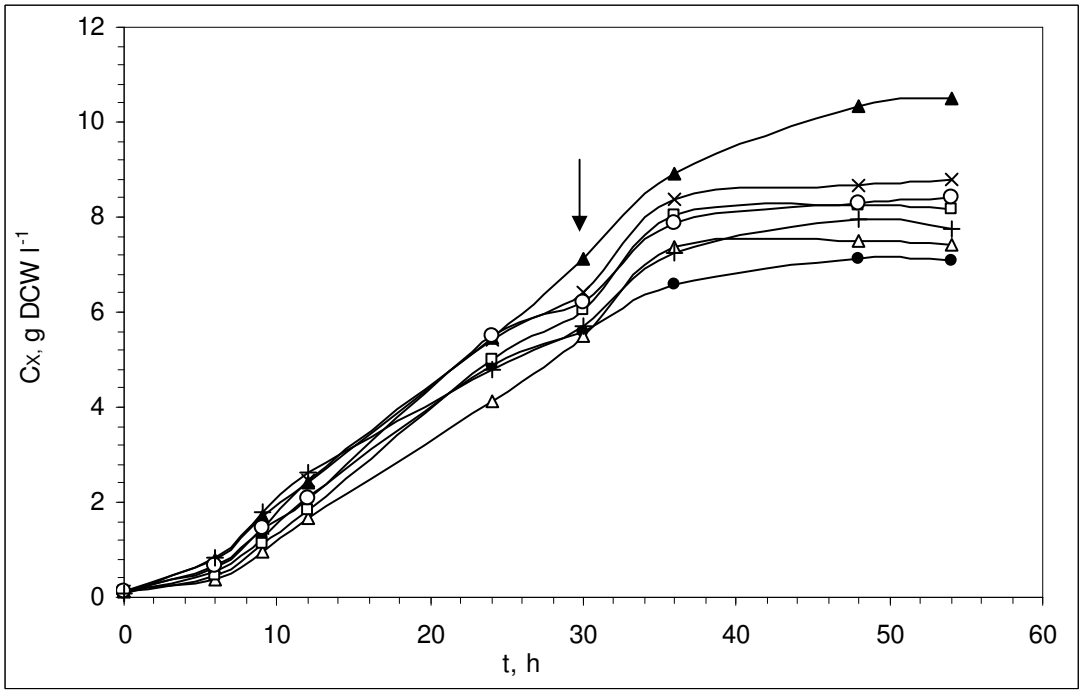


Figure 4.23 Variations in the cell concentration with the cultivation time in media containing different sources of glycerol: (●) pure glycerol (G1); (Δ) S1; (□) S2; (x) S3; (▲) C1; (O) C2; (+) C3. Arrow shows the point at which glycerol was totally consumed and methanol was added to the medium as the inducer of recombinant protein production.

Methanol, which is normally toxic, causes burden to the cells to be converted to non-toxic carbon source and consuming a lot of oxygen; though once converted, serves in the metabolic reactions. Functioning of the intracellular reaction network towards the byproduct formation, futile reactions or inefficient use of the carbon source would also explain the non-linear behavior of cell growth rate with substrate consumption rate. Moreover, the specific growth rates in the growth phase did not differ significantly from that obtained in the medium with pure glycerol (G1), $\mu = 0.21\ h^{-1}$; whereas in the induction phase, the growth rates were almost double the rate obtained in the medium with pure glycerol.

4.3.1.2.3 Effect of Biodiesel Byproduct Glycerol on Recombinant Protein Production

The effect of glycerol, present in the production medium and originating from different sources, on recombinant protein production of *P. pastoris* E17 was investigated by analyzing the supernatant of the medium, at the end of the process (t = 54 h). First, rHuEPO produced by *P. pastoris* cells was analyzed by dot blot analysis, a method which is fast and efficient to specifically detect a recombinant protein (Figure 4.24A); thereafter, the rHuEPO concentrations were quantified by HPCE (Figure 4.24B).

The intensity of the color, developed on the membrane blotted with the sample, is proportional with the concentration of the recombinant product (rHuEPO) within the sample (Figure 4.24A). The dot with the highest intensity was obtained for the medium with glycerol sample C1, which was significantly intense compared to the medium with pure glycerol. Thereafter, the numerical values for the rHuEPO concentration were obtained by capillary electrophoresis analysis, the results of which were in good correlation with the dot blot analysis (Figure 4.24B) and statistically different ($p<0.05$) for each source. The concentration of rHuEPO (C_P) in the medium with glycerol sample C1 (which gave the highest intensity dot in blot analysis) was the highest, $C_P = 31\ mg\ L^{-1}$, followed by S3 and C3. This was a consequential result since the media with glycerol samples C1 and S3, had the highest cell concentrations and with C3, the highest glycerol consumption rate was achieved. Thus, neither the cell concentration nor the glycerol utilization rates are directly proportional to the product formation rate,

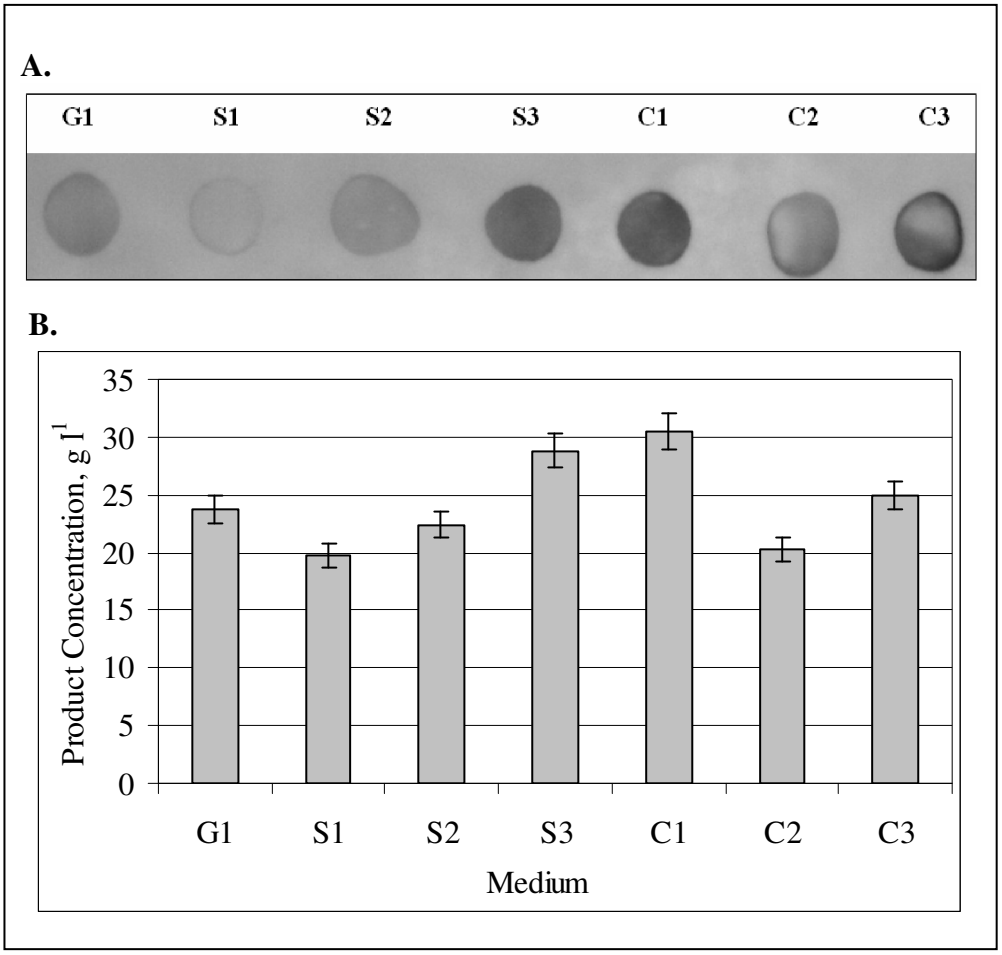


Figure 4.24 A. Dot blot analysis using 10 µl supernatant sample at t=54h of the bioprocess. Dot intensity is proportional only to the amount of rHuEPO produced in that sample. **B.** RHuEPO concentrations measured by HPCE in different media. The codes for different sources of glycerol used in the media are given in Table 4.1 and G1 stands for pure glycerol.

since the foreign gene product formation has in fact a complex relationship with many factors such as pH, oxygen supply, methanol uptake rate and extracellular protease activity. Moreover, since the glycerol concentration in the media was adjusted to the same concentration of 12.6 g L^{-1}, the concentration of methanol introduced at the same time varied for different glycerol sources. This could result in differences in cell growth and especially rHuEPO production; although the toxicity level of methanol in batch cultures (>3%) was not reached in any medium, at any time of the bioprocess. It is well known that glycerol represses AOX1 expression, however it is not known how methanol influences glycerol utilization and cell growth. On the other hand, the lowest product concentrations were obtained for the media with glycerol samples C2 and S1. Comparing rHuEPO production capacity of *P. pastoris* in medium with glycerol sample C1, to the medium with pure glycerol (G1), 1.3 times higher recombinant protein was obtained as seen from Figure 4.24B. Moreover, 1.5-fold difference observed between the highest (C1) and the lowest (S1) rHuEPO producing media demonstrates the importance of the proper selection of the glycerol source for the production medium.

4.3.1.2.4 Yields Coefficients

To better describe the bioprocess in more detail, stoichiometrically related parameters, namely yield coefficients, are defined. The overall yield of cell generated per mass of substrate consumed ($Y_{X/S}$), the overall yield of product formed per mass of substrate consumed ($Y_{P/S}$), and the overall yield of product formed per mass of cells generated ($Y_{P/X}$) were calculated for the rHuEPO production process (Table 4.2). Among the biodiesel byproducts, the highest $Y_{X/S}$ value was obtained with the glycerol sample C1, as $Y_{X/S}$=0.57 g dry cell weight g^{-1} substrate, followed by sample S3 having a value of $Y_{X/S}$= 0.51 g g^{-1}. These relatively higher cell yields, when compared to the medium with pure glycerol, points out the remarkable influence of the additional nutrients present in crude glycerol. The $Y_{P/X}$ value was the highest for samples C3 and G1 (pure glycerol) being statistically the same, followed by samples S3 and C1; having the values of $Y_{P/X}$=3.38, 3.35, 3.26 and 2.90 mg product g^{-1} dry cell weight, respectively. The highest $Y_{P/S}$ value was obtained with sample C1, having the value of 1.48 mg product g^{-1} substrate, followed by samples S3 and C3, which were significantly higher than that obtained with pure glycerol. Statistical difference of the mean values reported in Table 2 was verified, and $p<0.05$.

Table 4.2 Overall bioprocess yields attained for the rHuEPO production[a].

Sample Code	$Y_{X/S}$ (g cell g^{-1} substrate)	$Y_{P/X}$ (mg product g^{-1} CDW)	$Y_{P/S}$ (mg product g^{-1} substrate)
G1	0.44	3.35	1.16
S1	0.44	2.68	0.96
S2	0.48	2.72	1.08
S3	0.51	3.26	1.39
C1	0.57	2.90	1.48
C2	0.49	2.39	0.98
C3	0.45	3.38	1.27

a. Overall cell yield on substrate was calculated for the growth phase only (30h), thus considering only glycerol as the substrate. Other yields were calculated considering the whole process. Codes of the samples are as given in Table 4.1. G1 stands for pure glycerol.

For most yeast and fungi grown aerobically, $Y_{X/S}$ is typically between 0.4 - 0.8 g g^{-1}, having a value of $Y_{X/S}$ = 0.50 g g^{-1} for *S. cerevisiae* grown aerobically on glucose (Bailey and Ollis, 1986). These values (<1) indicate that the carbon source is used not only for biomass generation but also for energy generation, product, byproduct formation and maintenance. Thus, the values obtained for the cell yields on substrate are in a consistent range with the values found in literature. Considering the cell and product yields and intentions for choosing the favorable carbon source from various crude glycerol samples, S3, C1 and C3 are notable.

The yields of the bioprocess could further be improved by carrying out the production in controlled bioreactor systems, perhaps using growth models (Sinha et al., 2003; Zhang et al., 2005) to optimize the *P. pastori*s fermentation process. This will also eliminate the oxygen limitations rendered in laboratory scale shake bioreactor cultures, where oxygen, diffusing through the cotton air filters, is limiting to the cells by surface aeration. Future work to test for impurities and proteases that may be produced by changing the glycerol source would also be required for further process optimization.

4.3.2 Defined Medium Design

In the previous phases of this study, a common complex medium, BMMY, was sufficient for structural analysis, and thereafter an alternative complex medium was developed. However, industrial process prefer defined medium in cases where high purity is required in the downstream purification processes and where batch to batch variations resulting from complex media are intolerable. Thus, for the production of rHuEPO, a defined medium is more preferable, as well as for bioprocess modeling studies in general.

The search for a defined production medium, optimum for the cell growth of recombinant *P. pastoris*, was carried out by screening the most commonly used salt solutions in literature for *P. pastoris*. This strategy was followed in order to expand the usability of the models that will be formed subsequently. Thereafter, adding sorbitol batch wise to the induction phase of *P. pastoris* was examined and in the medium containing sorbitol, methanol induction strategies were investigated. Moreover, the non-inhibitory sorbitol concentration was determined for further use in high-cell density pilot scale bioreactor processes.

4.3.2.1 Effects Macro and Micronutrients on Recombinant *P. pastoris* Growth

For the purpose of designing a defined medium, the effects of macro and micronutrients were investigated first. Rather than screening several salts individually, the most commonly used salt solutions in literature for *P. pastoris* were screened for optimizing recombinant *P. pastoris* growth.

The nitrogen and the carbon sources in the production medium were fixed as 7.9 kg m^{-3} of methanol fed every 24 h and 10 kg m^{-3} of $(NH_4)_2SO_4$. All the media were buffered with 0.1 M phosphate buffer at pH=6.0. The differences in macronutrients and trace elements are given in Table E.1 and E.2 (Appendix E) respectively, together with their references. These different macro and micronutrient solutions were screened in various combinations to optimize cellular growth. The experimental design of this combination is given in Table 4.3.

Table 4.3 Experimental design for investigating the combined effect of salts on cellular growth[a].

Medium #	Macronutrient solution	Micronutrient solution
1	BSM	PTM1
2	BSM	PTM4
3	mBSM	PTM1
4	mBSM	PTM4
5	m2BSM	PTM4
6	m3BSM	PTMJ

a. The abbreviated solutions are given in Table E.1 and E.2

For studying the effect of various macro and micronutrients on rHuEPO production by recombinant *P. pastoris*, it was sufficient to investigate the effects on cell concentration. Thus, by simple logic, more cells generally mean more product. Although the high concentration of calcium in BSM + PTM1 medium results in high amount of precipitation when combined with other components of the medium, it is the most commonly used medium in the literature. This issue has been addressed many times in the literature, and Brady et al. (2001) found that decreasing the salt content (i.e. using the mBSM medium) reduced the formation of lipid-like substances during the course of fermentation, compared to the standard salts medium, i.e., BSM+PTM1. However, in the medium containing BSM+PTM1, highest cell concentration was obtained, C_X=7.2 kg m^{-3}, although close to the mBSM+PTM1 salt solution (Figure 4.25), higher calcium and magnesium content of BSM compared to mBSM would be necessary in high-cell density cultivations. In lower concentrations of trace elements (PTM4 and PTMJ), cellular growth was observed to be insufficient (Figure 4.25). In conclusion, the defined medium containing BSM+PTM1 salt solution was found to be the optimum condition for cell generation and will be used in further studies.

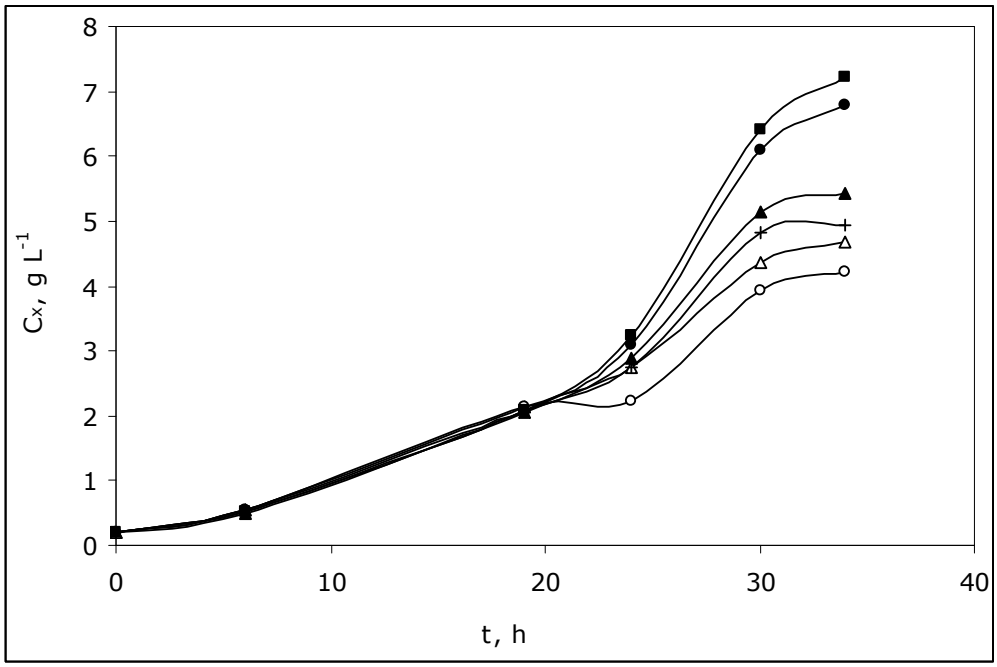


Figure 4.25 Effects of macro and micronutrients on cell concentration. The media contained: BSM+PTM1 (■); mBSM+PTM1 (●); BSM+PTM4 (▲); mBSM+PTM4 (Δ); m2BSM+PTM4 (O); m3BSM+PTMJ (+).

4.3.2.2 Effect of Initial Sorbitol Concentration on Recombinant *P. pastoris* Growth and RHuEPO Production

Sorbitol is a non-repressing carbon source for *AOX* promoter, and although widely utilized for Mut^S strains (methanol utilization slow phenotype, where only *aox2* is intact, which is responsible for 15% of the AOX activity), have recently been appreciated for Mut^+ strains (wild type *AOX*, where both alcohol oxidase genes *aox1* and *aox2* are intact), and investigated in only a limited number of studies (Sreekrisha et al., 1997; Inan et al., 2001-b; Jungo et al., 2007; Ramon et al., 2007). However, these studies in the literature have not fully investigated the batch or fed-batch operational strategies in the presence of sorbitol nor the effect of sorbitol on cell metabolism of the Mut^+ *P. pastoris* strain.

Thus, with the aim of incorporating sorbitol to the bioprocess by Mut^+ *P. pastoris* strain, firstly the effect of initial sorbitol concentration on cell growth and rHuEPO production was investigated. The BSM+PTM1 production media, induced by methanol only at the beginning of the process, contained different initial sorbitol

concentrations, C_{S0} = 0, 2, 5, 7.5, 11.5 and 18 g L^{-1}, which corresponds roughly to %C-mol sorbitol of 0, 20, 40, 50, 60 and 70%.

With the increasing amounts of initial sorbitol concentrations, cell concentration increased as expected (Figure 4.26). Moreover, recombinant protein production profiles were examined from dot-blot analysis (Figure 4.27), where the dot intensity is only proportional to the amount of recombinant protein produced. A general trend of increase in rHuEPO concentration with increasing initial sorbitol concentration until C_{S0}= 7.5 g L^{-1} was observed. Another interesting observation from Figure 4.27 was that at t=24 h, where cell concentrations in all media were almost the same (Figure 4.26), the rHuEPO concentration was generally increasing with increasing initial sorbitol concentration. Thus, the positive effect of sorbitol on recombinant protein production was not only due to increase in cell concentration, but likely also due to increased secretion of intracellular recombinant protein. Considering that production profiles reached its maximum at C_{S0}= 7.5 g L^{-1}, this concentration was found to be sufficient for laboratory scale shake-bioreactors.

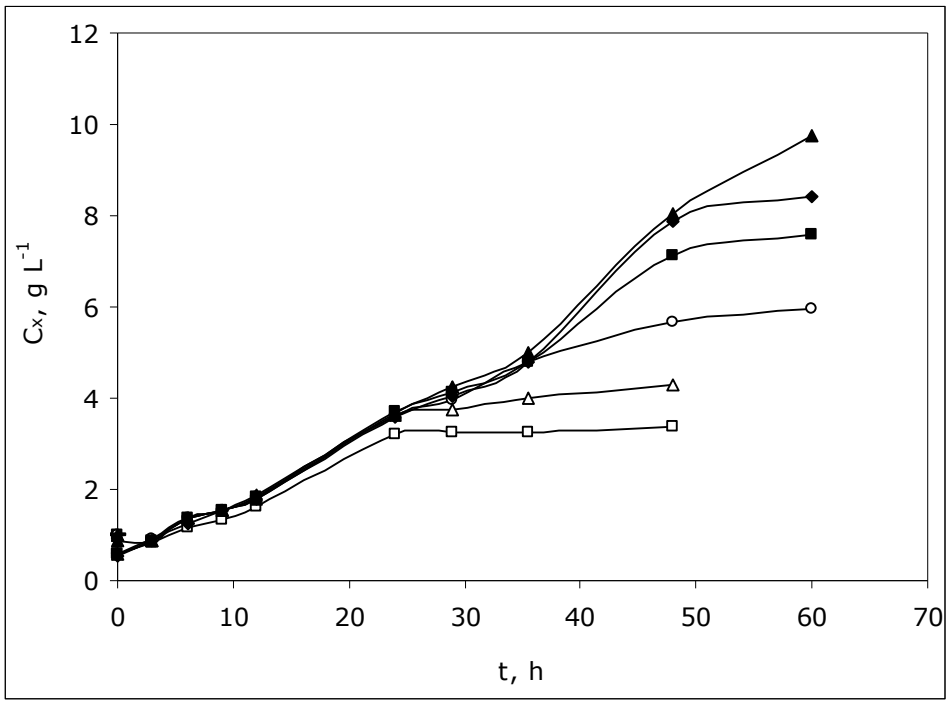


Figure 4.26 Variation in cell concentration in BSM+PTM1 media, induced with 1% methanol at t=0h only, and containing different initial sorbitol concentrations, C_{S0}(g L^{-1}): 0, (□); 2, (Δ); 5, (O); 7.5, (■); 11.5 (♦); and 18, (▲).

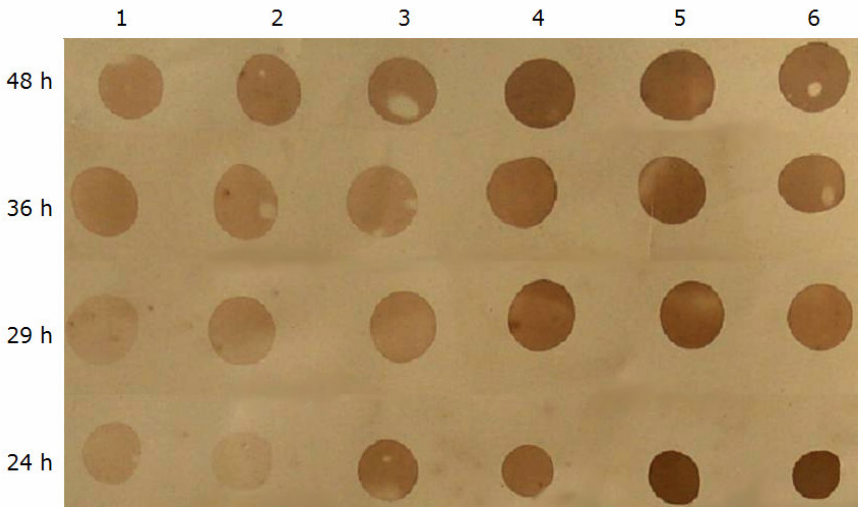


Figure 4.27 Dot blot analysis view of 10 µl supernatant samples taken at different cultivation times of the bioprocess, from media containing different initial sorbitol concentrations, C_{S0} (g L^{-1}): Lane 1: 0; Lane 2: 2; Lane 3: 5; Lane 4: 7.5; Lane 5: 11.5 and Lane 6: 18.

4.3.2.3 Effect of Methanol Induction Strategy on Recombinant *P. pastoris* Growth

In view of the fact that adding sorbitol batch wise to the production medium at an initial concentration of C_{S0}= 7.5 g L^{-1} was found beneficial for cellular growth and recombinant protein production in laboratory scale shake bioreactors, where methanol was given to the medium only t=0h, the methanol induction strategy was studied next, in BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol. It was found that inducing with 1% (v/v) methanol at t=24 h was found to be crucial, perhaps at the point where methanol was totally consumed (Figure 4.28), and showing a more desirable growth profile than that attained in the medium induced at t=36 h. Thereafter, the effect of methanol concentration used for induction on cell growth was investigated in the BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol and in the medium induced with 1% (v/v) methanol, the highest cell concentration was attained, C_X=9.1 g L^{-1} (Figure 4.29).

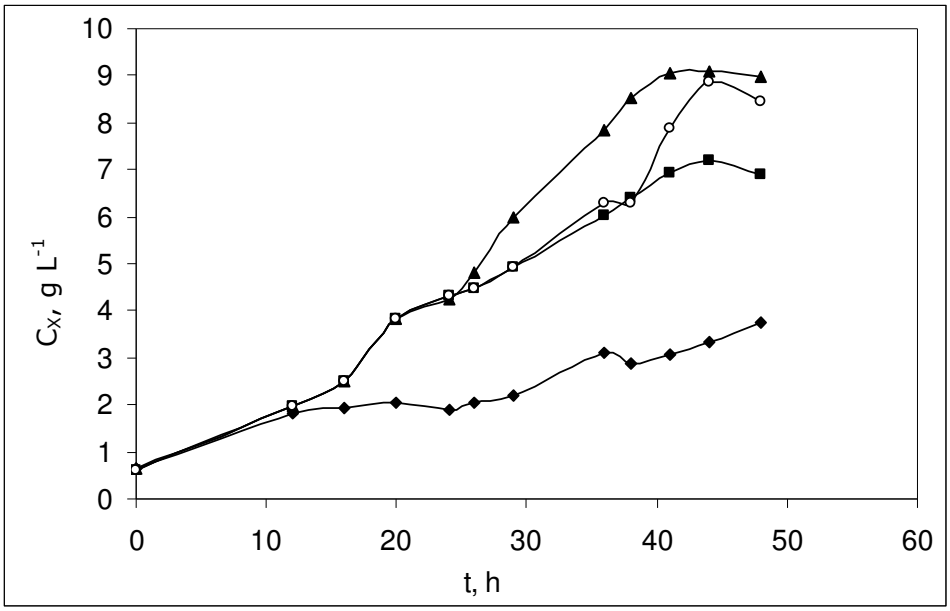


Figure 4.28 Variation in cell concentration with cultivation time in BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol and different methanol induction strategy. (♦): no methanol induction; (■): induction at t=0 h with 1% (v/v) methanol; (o): induction at t=0 and 36 h with 1% (v/v) methanol; (▲): induction at t=0 and 24 h with 1% (v/v) methanol.

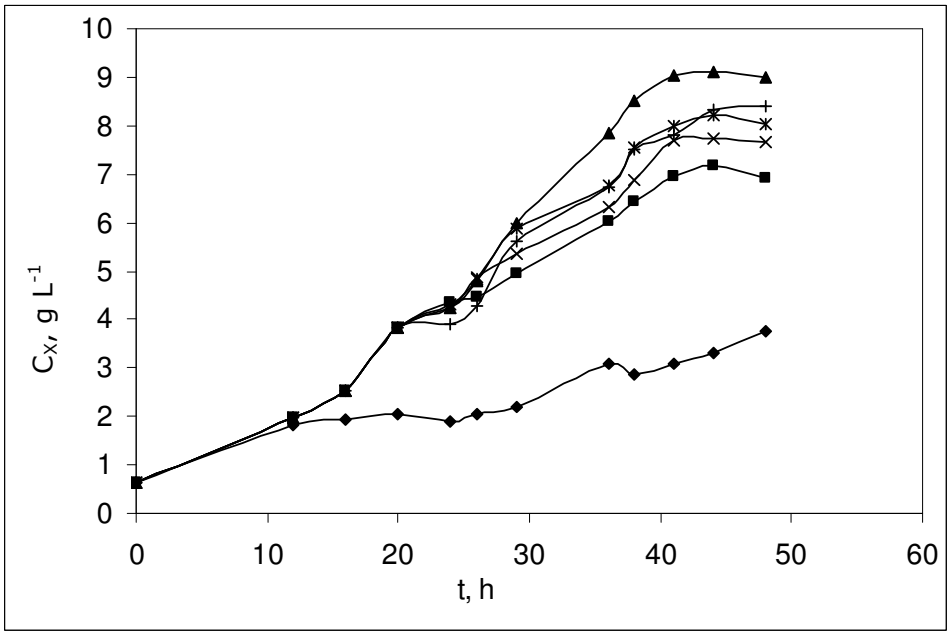


Figure 4.29 Variation in cell concentration with cultivation time in BSM+PTM1 medium containing 7.5 g L^{-1} sorbitol and different methanol induction strategy. (♦): no methanol induction; (■): induction at t=0 h with 1% (v/v) methanol; (x): induction at t=0 and 24 h with 0.25% (v/v) methanol; (∗): induction at t=0 and 24 h with 0.5% (v/v) methanol; (+): induction at t=0 and 24 h with 0.75% (v/v) methanol; (▲): induction at t=0 and 24 h with 1% (v/v) methanol.

4.3.2.4 The non-Inhibitory Sorbitol Concentration

As sorbitol was found as an advantageous carbon source in many aspects and decided to be added to the fermentation medium batch-wise during the production phase, the non-inhibitory concentration of sorbitol had to be investigated to be used in high-cell density fed-batch pilot scale bioreactor operations, where the initial sorbitol concentration optimized for laboratory scale shake bioreactors (C_{S0}=7.5 g L^{-1}) would be insufficient. With the addition of another carbon source, the amount of cells generated is expected to increase until an inhibitory level of this new substrate is reached. As more cells simply mean more recombinant product and literature has no reports on decreased product yields in sorbitol supplemented medium, there was no need to investigate the effect of sorbitol on product formation, at this stage.

In 50 mL defined production medium (BSM+PTM1), the effect of initial batch sorbitol concentration on *P. pastoris* growth was investigated, and it was found that the highest cell concentration (C_X= 10 kg m^{-3}) was reached in the medium supplemented with 50 g L^{-1} sorbitol, which was 1.5-fold higher than that obtained in medium containing methanol as the sole carbon source (Figure 4.30). Thus, 50 g L^{-1} was found as the non-inhibiting concentration limit for sorbitol (Figure 4.31).

4.3.3 Bioreactor Operation Parameters

When designing a bioprocess, determining the ideal bioprocess operation parameters (pH, temperature and dissolved oxygen level) is important to obtain high productivity and yields. In this study, where the aim is to design and mathematically model the bioprocess for recombinant human erythropoietin by *Pichia pastoris* cells, the temperature has been set at 30°C. This is the most frequently applied condition for *Pichia pastoris* processes, regardless of the product. Although depending on the product, lowering the temperature was reported to potentially increase the protein yields due to lower proteolytic activities (Daly and Hearn; 2005), the optimum growth temperature, T=30°C, for the microorganism was chosen for modeling purposes. On the other hand, since dissolved oxygen level can not be controlled for laboratory scale shake-bioreactors, N=250 rpm was chosen.

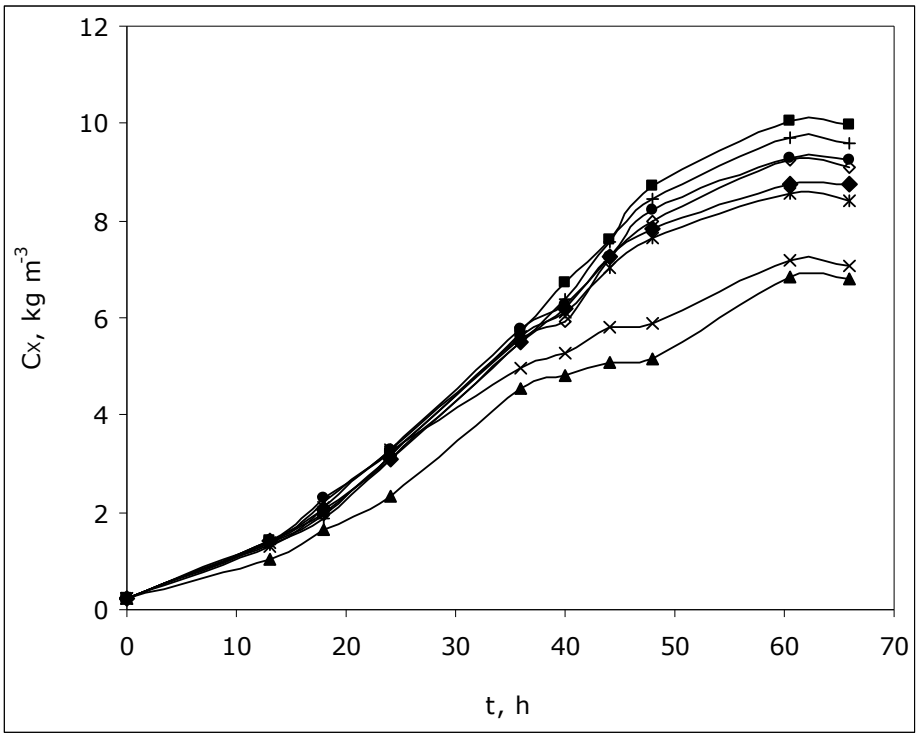


Figure 4.30 The variation in cell concentration with cultivation time in media with different initial sorbitol concentrations, C_{S0} (kg m^{-3}): 0, (▲); 2, (x); 10 (∗); 25, (●); 40, (+); 50, (■); 60, (◇); 80, (♦).

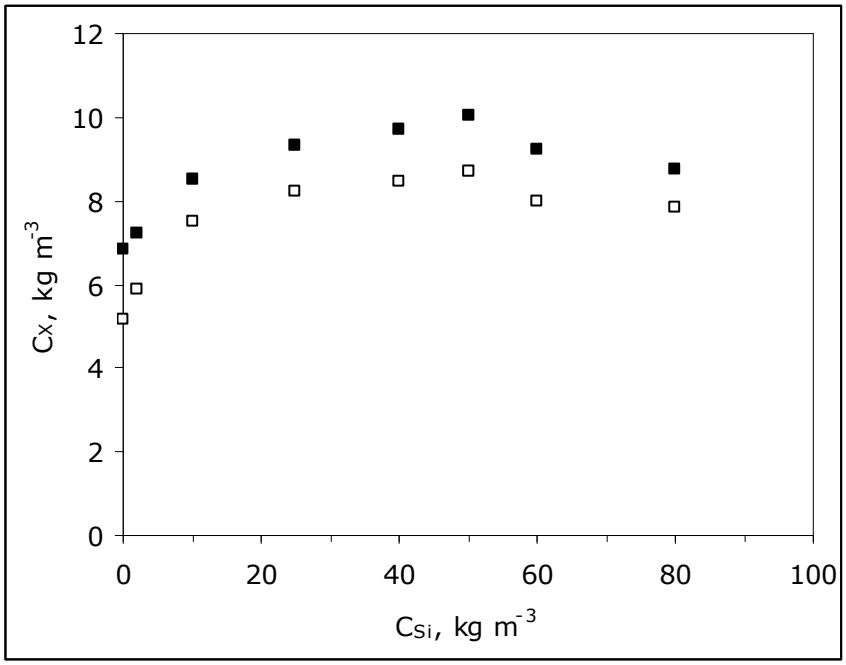


Figure 4.31 Variation in cell concentration with initial sorbitol concentration, at t=48 h (□) and t=60h (■) of the bioprocess.

However, pH has to be optimized for each bioprocess, as pH has a diverse effect on each specific protein, influencing proteolysis and stability of the protein. Although cells have a remarkable ability to maintain their intracellular pH at a constant level at the expense of a significant increase in the maintenance demands, rHuEPO, which was designed to be an extracellular protein, would be affected by the medium pH. While some bioprocesses require controlled pH conditions, others might require uncontrolled pH operations. Moreover, *P. pastoris* is capable of growing across a relatively broad pH range (3.0–7.0) (Macauley-Patrick et al., 2005). Thus, in this study, the effect of pH on cell growth and recombinant protein production was investigated in laboratory scale shake bioreactor experiments. Moreover, the effect of medium pH on the stability of rHuEPO produced in bioreactor experiments was also investigated.

4.3.3.1 Effect of pH on RHuEPO Production and Stability

Firstly, the effect of medium pH on cell growth and recombinant protein production was investigated in laboratory scale shake bioreactor experiments. It was observed that the pH=5.0-6.0 range were the mostly favored conditions (Figure 4.32).

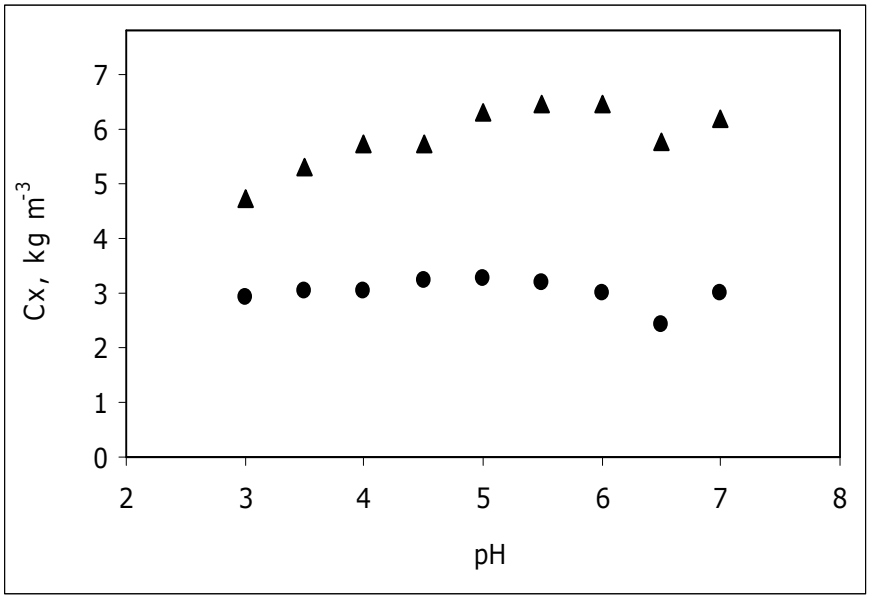


Figure 4.32 Variation in cell concentration with respect to medium pH, at t= 24 h (●) and t=43 h (▲) of the bioprocess.

On the other hand, in terms of recombinant protein productivity, in the production media containing the pH=4.0-5.0 buffers, the higher product concentrations are achieved, while after pH=5.0, the productivity decreases (Figure 4.33). Thus, pH =5.0 is the optimum pH, which differs from the pH of complex medium used up to this point, pH=6.0. Although *P. pastoris* has been reported to be capable of growing across a relatively broad pH range (3.0–7.0) (Macauley-Patrick et al., 2005), the most commonly applied conditions are pH controlled conditions, in the range of pH=5-6. Nevertheless, the results obtained may have been slightly affected by the fact that the final pH measured in all the media was around pH=2.2, thus the buffering capacity of the citric acid-phosphate buffer was not enough for the 43 h bioprocess.

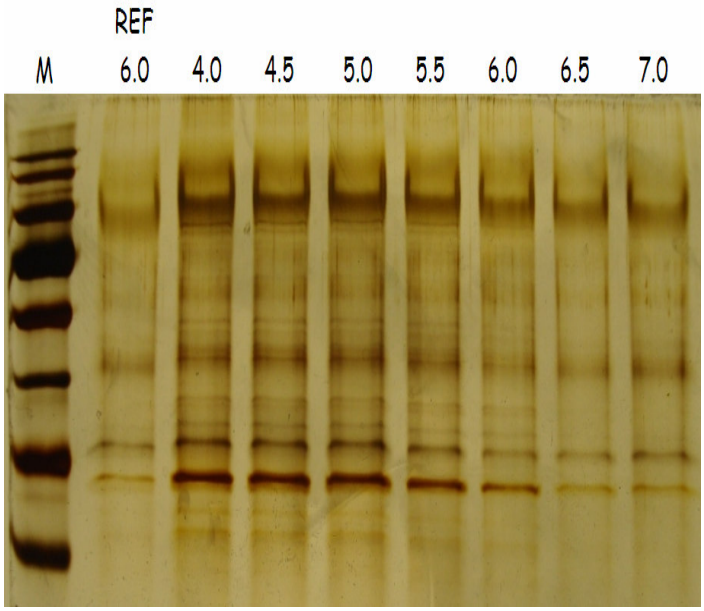


Figure 4.33 Silver stained SDS-PAGE gel view of extracellular proteins produced by *Pichia pastoris* in shake-bioreactor experiment, to observe the effect of medium pH on rHuEPO production. **M:** Pre-stained Protein Ladder. The numbers above the gel indicate the pH of the extracellular medium.

4.4 Expression of Recombinant Human Erythropoietin by *P. pastoris* in Pilot Scale Bioreactors

Having developed the recombinant microorganism producing rHuEPO, purified and characterized the recombinant protein produced by the microorganism, and having investigated the effect of medium components and pH on rHuEPO production and cell growth in laboratory scale air filtered shake bioreactors, the next step in the bioprocess development was to scale up to a more realistic, pilot-scale bioreactor, which is equipped with temperature, pH, foam, stirring rate, feed inlet rate and dissolved oxygen control. After implementing mass flow controllers to the bioreactor system in order to fine tune the dissolved oxygen concentration in the medium, four sets of fed-batch bioreactor experiments were performed, to investigate the effect of feeding rate of the carbon source, methanol, and the effectiveness of the proposed production strategy of adding sorbitol in batch-mode at the induction phase as a co-substrate.

4.4.1 Control of Bioreactor Operation Parameters in Pilot-Scale Bioreactor

Temperature, pH, dissolved oxygen level, and agitation rate are the bioreactor operation conditions that have to be controlled as accurately as possible in bioprocesses.

The temperature was accurately controlled at T=30±0.1°C by the PI controller of the bioreactor system. For pH control, first 25% NH_3OH and 30% H_2SO_4 was tried but the pH control was not successful and there was too much volume increase due to H_2SO_4 solution. Then it was found that using only the base (25% NH_3OH) and using the PI control parameters as Xp=30% and TI= 30s, and keeping the base-pump-valve open at 10%, the pH of the medium was automatically controlled at pH=5.0 ± 0.3. Therefore, H_2SO_4 solution was not used in the experiments for pH control. Foaming was prevented by adding antifoam solution to the initial medium (0.01% v/v), and was never required thereafter, as the working volume was below 1.0 dm^3 and aeration was provided from a sparger at the bottom of the vessel. For agitation rate, N=750 rpm was tried in initial experiments, which is considered a high level for bacterial fermentations. However, this was not enough for *P. pastoris* yeast cells, which

consumes oxygen at a very high rate. Thus, an agitation rate of N=900 rpm was used instead, and kept constant throughout the cultivation period. A higher agitation rate can cause shear damaging the cells, increase in temperature and foam formation.

The dissolved oxygen level was the most problematic one to control. As *Pichia pastoris* has extremely high oxygen consumption rates and for oxygen not to be a limiting factor, usually the aim is to maintain its level above 20% saturation. After 12 h of fermentation, it was not possible to keep the dissolved oxygen (DO) levels above 20% saturation, using air. Therefore, instead of air, firstly pure oxygen was given to the system after t=12h. However, the DO% values fluctuated too much although the PID control parameters were tried to be optimized (Figure 4.34). Thus, air had to be enriched with increasing amounts of oxygen. For this purpose, mass flow controllers were integrated to the oxygen and air lines, prior to where the lines meet before entering the bioreactor system (Figure 3.1). As a result, the fluctuation in DO% level was reduced about 5-fold (Figure 4.34).

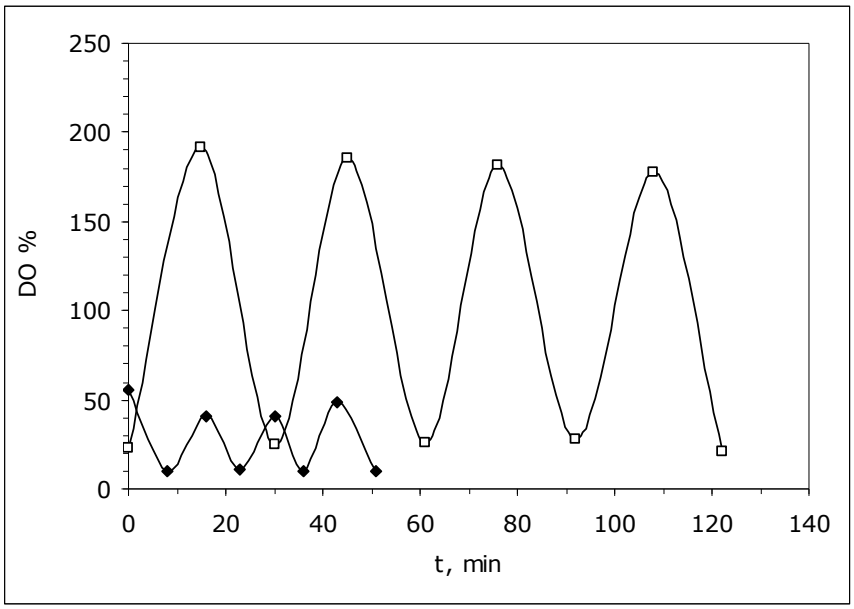


Figure 4.34 A screenshot of the dissolved oxygen levels during two bioreactor operations, where controls were by pure oxygen (□), or oxygen enriched air (♦).

4.4.2 Effect of Sorbitol and Methanol Feeding Rate in Fed-Batch Pilot Scale Bioreactor Operations

Specific growth rate being the most important variable in fed-batch fermentations, determines the exponential feed inlet rate of the limiting substrate methanol, according to equation 4.1 (Weigand et al., 1979),

$$F(t) = \frac{\mu_0 V_0 C_{X0}}{C_{Mo} Y_{X/M}} \exp(\mu_0 t) \qquad (4.1)$$

where, μ_0 is the desired specific growth rate, V_0 is the initial volume, C_{X0} is the initial cell concentration, C_{Mo} is feed substrate (methanol) concentration and $Y_{X/M}$ is the cell yield on methanol. The chosen defined medium, BSM+PTM1, was supplemented with 50 g L^{-1} sorbitol, which is added to the medium in the induction phase, and effect of methanol feed rate was investigated. Firstly, for μ_0=0.03 h^{-1}, production was carried out in the reference medium, BSM+PTM1 without the sorbitol. The specific growth rate (μ_0=0.03 h^{-1}) was selected for a number of reasons. Firstly, Cunha et al. (2004) found that above specific growth rates of μ=0.025 h^{-1}, specific recombinant protein productivity does not depend on the growth rate. Moreover, the maximum specific growth rate of *Pichia pastoris* on methanol is μ_{max}=0.14 h^{-1}, while that on sorbitol is μ_{max}=0.032 h^{-1} (Jungo et al., 2007), thus the specific growth rate was intended to be kept close to the maximum specific growth rate on sorbitol. Moreover, high specific growth rates (μ>0.038 h^{-1}) were shown to promote glycosylation (Schenk et al., 2008), which is an unwanted situation.

Then, using this predetermined methanol feed profile for μ_0=0.03 h^{-1}, 50 g L^{-1} sorbitol was added to the medium in batch mode. In the two following experiments, the predetermined feed profiles of methanol for μ_0=0.02 h^{-1} and μ_0=0.04 h^{-1} were used, where 50 g L^{-1} sorbitol was again present in the medium. The experimental design is explained in Table 4.4, giving each process a code for the ease of later data analysis. The predetermined feed profiles calculated according to equation 4.1 are given Figure 4.35. The parameters used in equation 4.1 are given in Table 4.5. The feed inlet rate was controlled automatically by the bioreactor control unit and verified by hourly taken data with a balance placed under the methanol feed bottle.

Table 4.4 The experimental plan for bioreactor experiments

Experiment Code	Predetermined Specific Growth Rate, μ, h^{-1}	Presence of Sorbitol (Batch, 50 g L^{-1})
M-0.03	0.03	No
MS-0.03	0.03	Yes
MS-0.02	0.02	Yes
MS-0.04	0.04	Yes

Table 4.5 Parameters used in equation 4.1

Parameter	Glycerol Fed-Batch Phase	Methanol Fed-Batch Phase
μ_0 (h^{-1})	0.18	0.02 or 0.03 or 0.04
$Y_{X/M}$ (g g^{-1})	0.5 *	0.4 **
C_{Mo}(g L^{-1})	1130	790

* Ref: Cos et al., 2005

** Ref: Jungo et al., 2006

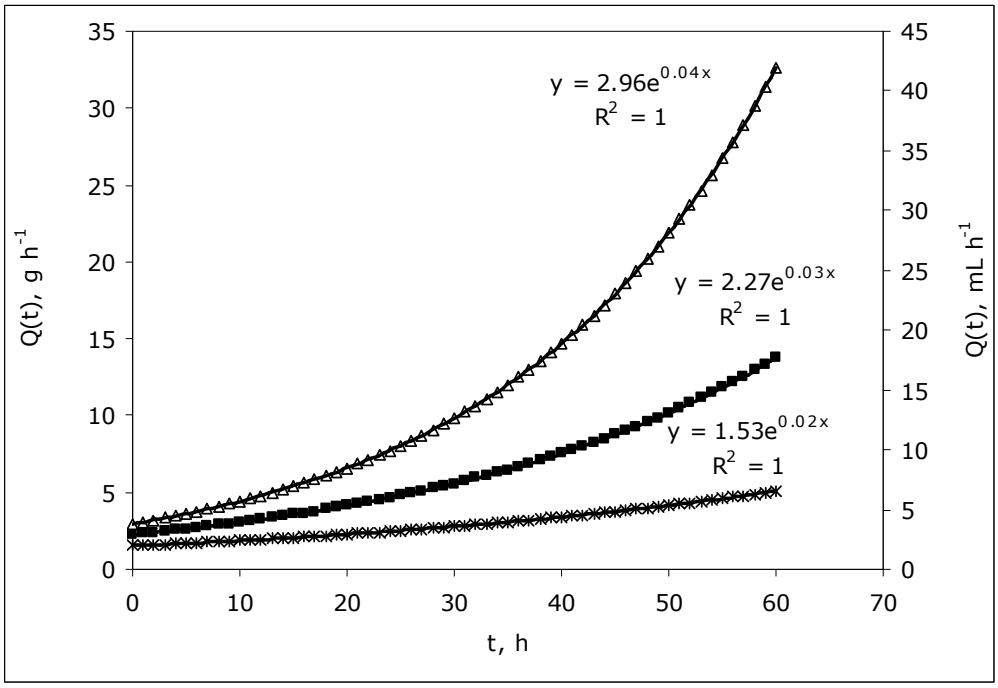


Figure 4.35 The predetermined feed profiles for methanol, calculated for each specific growth rate (μ_o) aimed, 0.02 h^{-1}, (∗); 0.03 h^{-1}, (■); 0.04 h^{-1}, (Δ). The profile equations are given in g h^{-1} units.

4.4.2.1 Effect of Sorbitol and Feeding Rate of Methanol on Cell Growth

The most explicit effect of medium design and feeding strategy should be examined firstly on cell growth. The cell growths attained in the precultivation phase of all runs were nearly the same (Figure 4.36a). Although the process is continuous and the production (induction) phase starts at around t=32h of the precultivation period, for simplicity in calculations, commence of the production phase was taken as t=0h (Figure 4.36b).

Adding sorbitol batch wise to the medium, was beneficial in terms of eliminating the long lag-phase (t=0-9 h) for the cells and therefore cell yields were greatly improved (Figure 4.36b). For instance at t=24 h, 1.7-fold higher cell concentration was obtained in medium with sorbitol (MS-0.03) compared to that without sorbitol (M-0.03). Moreover, comparing MS-0.03 and M-0.03 media, although the same methanol feeding profile was used, the growth rate was improved in the presence of sorbitol as expected. Thus, the cells are protected from the negative effects of methanol on cell metabolism, and can start proliferating immediately in the induction medium.

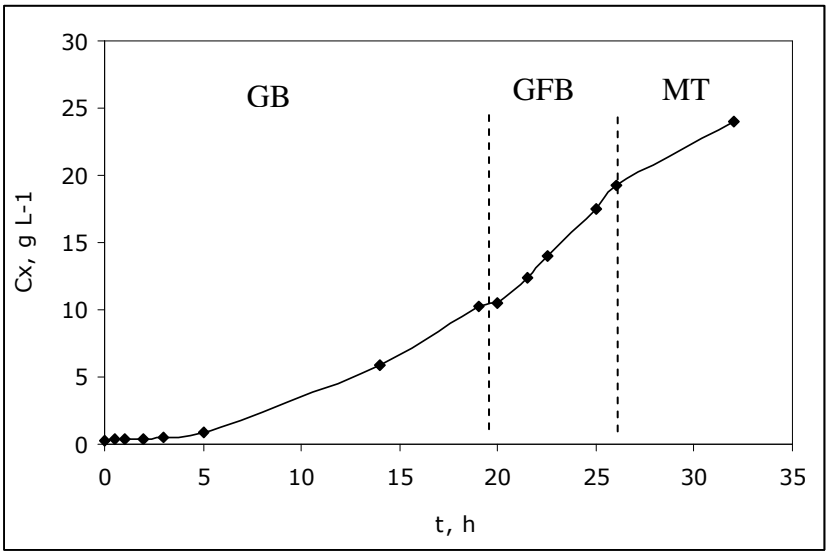


Figure 4.36-a. Variation in cell concentration with cultivation time in the precultivation phases: glycerol batch phase (GB), glycerol fed-batch phase (GFB) and methanol transition phase (MT).

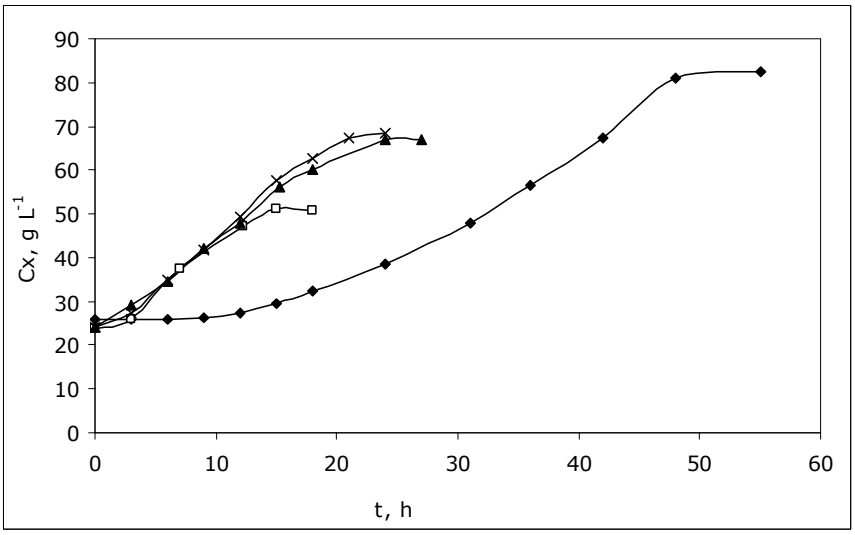


Figure 4.36-b. Variation in cell concentration with cultivation time in the production phase of bioprocesses with different feeding profiles: M-0.03 (♦); MS-0.02 (□); MS-0.03 (▲); MS-0.04 (x).

4.4.2.2 Effect of Sorbitol and Feeding Rate of Methanol on Recombinant Protein Production

The most important aim of bioprocess optimization studies is to increase the recombinant protein production. The effect sorbitol on rHuEPO production was directly observed by SDS-PAGE analysis, where the products obtained at the same cultivation time were compared (Figure 4.37) ; and then measured by HPCE and Bradford assay (Figure 4.38). Thus, 1.8-fold higher rHuEPO was obtained at t=18 h in MS-0.04 condition compared to M-0.03 condition. Moreover, the methanol feeding rate influenced productivity as expected, since *AOX1* promoter is induced by methanol, as clearly seen when MS-0.02 and MS-0.03 media are compared in Figure 4.37 (Lanes 2 and 3). On the other hand, rHuEPO concentration detected in the extracellular media from MS-0.03 and MS-0.04 experiments were nearly the same, yielding 130 mg L^{-1} rHuEPO.

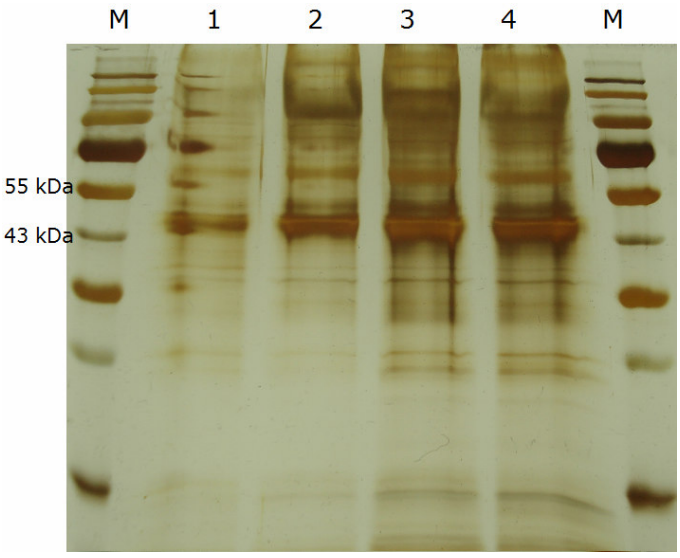


Figure 4.37 SDS-PAGE gel view of extracellular production medium proteins from bioprocesses with different feeding profiles. M: Protein ladder, Lane 1: M-0.03, t=24 h; Lane 2: MS-0.02, t=18 h; Lane 3: MS-0.03, t=24 h; Lane 4: MS-0.04, t=24 h.

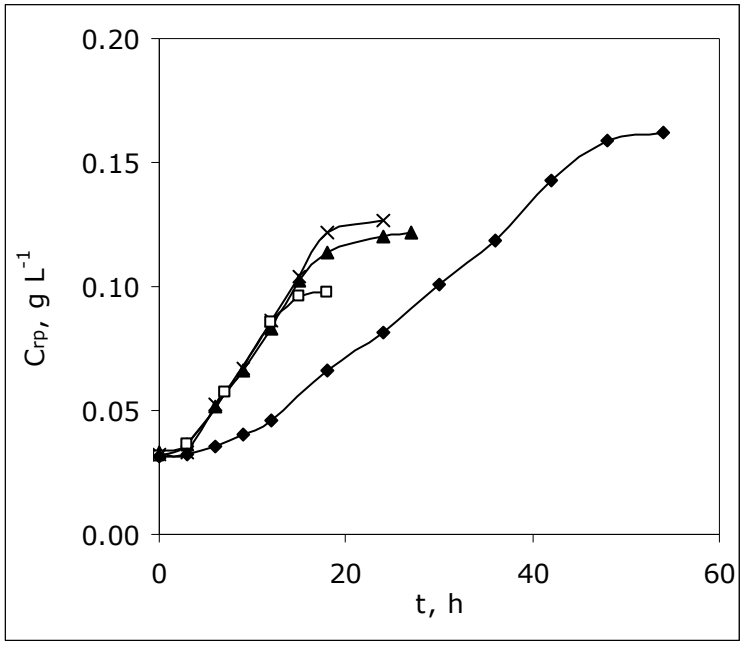


Figure 4.38 Variation in recombinant protein concentration with cultivation time in the production phase of bioprocesses with different feeding profiles: M-0.03 (♦); MS-0.02 (□); MS-0.03 (▲); MS-0.04 (x).

4.4.2.3 Effect of Methanol Feeding Rate on Sorbitol Consumption Rate

Having clearly observed the positive effects of sorbitol on cell and recombinant protein production, the effect of methanol feeding rate on sorbitol consumption rate was investigated. The two important findings (Figure 4.39) were that methanol feeding rate did not affect sorbitol consumption rate and that sorbitol utilization started immediately, and lasted for 18 h. As methanol was never detected in the medium, it was assumed that the cell consumes all methanol fed, with the feeding profile given in Figure 4.35. This means simultaneous utilization of sorbitol with methanol, confirming the results of Jungo et al. (2007) rather than Ramon et al. (2007), who claimed sequential utilization of sorbitol and methanol.

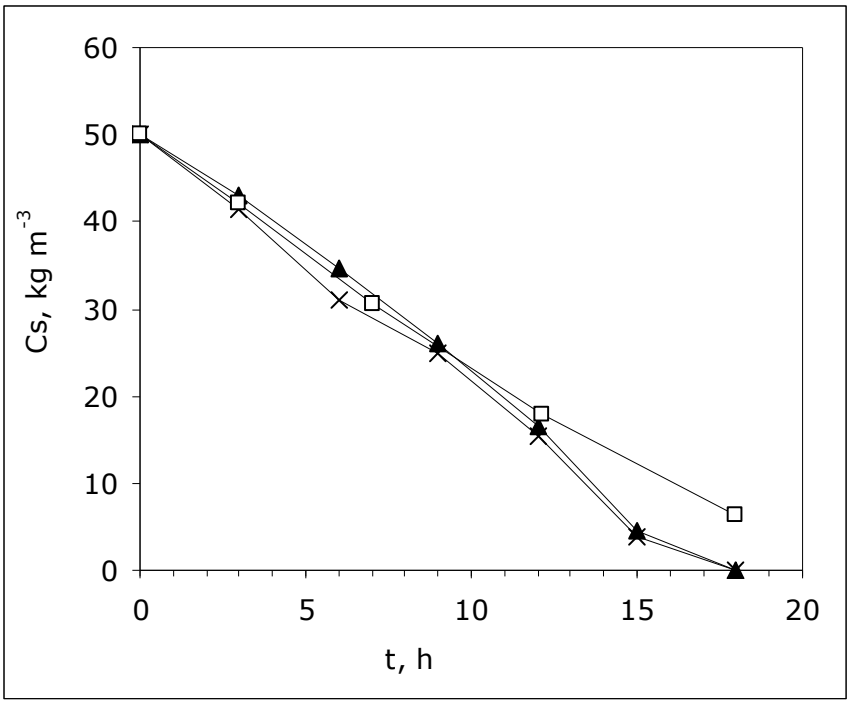


Figure 4.39. Variation in sorbitol concentration with cultivation time for different feeding profiles: MS-0.02 (□); MS-0.03 (▲); MS-0.04 (x).

4.4.2.4 Effect of Sorbitol and Feeding Rate of Methanol on Total Protease Production

The protease activities of acidic, neutral and alkali proteases in the *Pichia pastoris* production medium were measured separately and their total activity was then converted to total protease concentration. It was observed that as the methanol feed rate increased, the total protease concentration increased, reaching a maximum of C_{pro}=0.25 g L^{-1} in MS-0.04 condition (Figure 4.40). The obvious reason is that more cells would produce more proteases. However, when the specific protease concentrations are compared at t=24 h, C'_{pro}=2.60 mg protease g^{-1} cell dry weight for MS-0.03 condition is much less than C'_{pro}=3.06 mg protease g^{-1} cell dry weight for M-0.03 condition. Thus, it can be concluded that presence of sorbitol reduces protease production.

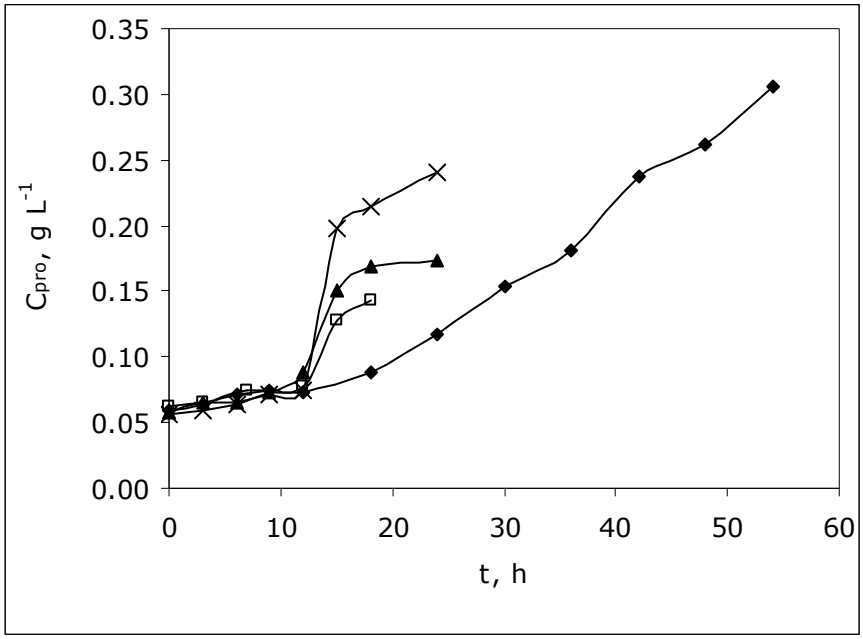


Figure 4.40. Variation in total protease concentration with cultivation time for different feeding profiles: M-0.03 (♦); MS-0.02 (□); MS-0.03 (▲); MS-0.04 (x).

Moreover, it was observed that while the neutral proteases were almost the same for all conditions, twice as much basic protease was observed in MS-0.04 condition compared to other conditions. Thus, basic proteases were affected the most by the feeding rate of methanol.

4.4.2.5 Effect of Sorbitol and Feeding Rate of Methanol on Alcohol Oxidase Production

RHuEPO production is coupled to alcohol oxidase (AOX) production, which is induced by the presence of methanol. Therefore, the level of intracellular AOX or relevantly its activity, should demonstrate rHuEPO translation levels as well. The specific AOX activity results showed a linear relationship to increasing feeding rates of methanol, thus at MS-0.04 condition the highest specific activity was obtained as 82 U g^{-1} CDW and the highest specific activities at M-0.03 and MS-0.03 conditions were the same at t=6 h (Figure 4.41). The enzyme had its highest activity at t=6h of the bioprocess in all conditions, however slight inhibition by sorbitol was observed, contrary to the findings of other studies.

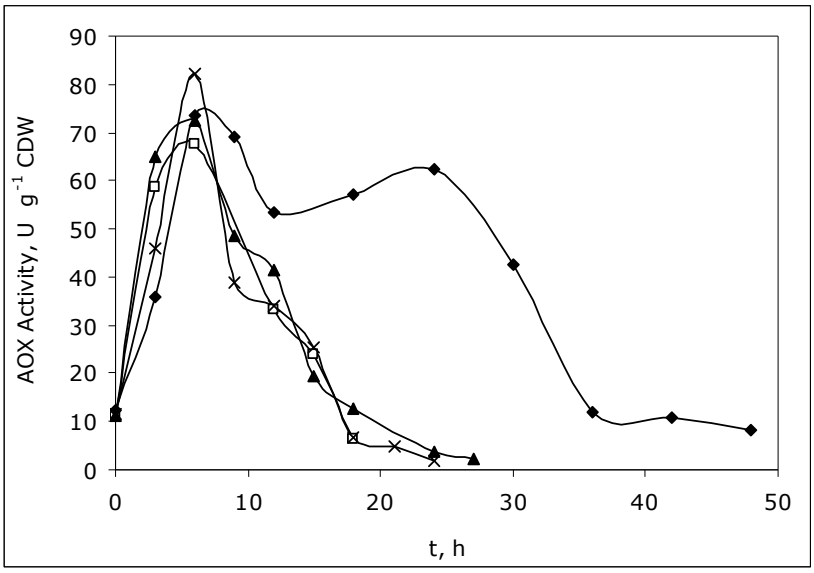


Figure 4.41 Variation in alcohol oxidase activity with cultivation time for different feeding profiles: M-0.03 (♦); MS-0,02 (□); MS-0,03 (▲); MS-0,04 (x).

4.4.2.6 Effect of Sorbitol and Feeding Rate of Methanol on Amino and Organic Acid Concentration Profiles

The variations in amino acid and organic acid concentrations in the fermentation broth with cultivation time are given in Table 4.6 and Table 4.7 respectively. Moreover, the results for lactic acid, which is also an organic acid, is given separately in Figure 4.42, since it is produced in largest amount. The amino and organic acid profiles give an idea about the supply and demand for these metabolites, which is regulated by the metabolic reaction network.

For the production of proteins, most microorganisms have the metabolic machinery to synthesize all essential amino acids from carbon and nitrogen sources; and regulation of the metabolic reaction network should cause good coupling of supply and demand for the amino acids. If the supply exceeds the demand, the amino acids would be excreted to the extracellular medium and if they are required again, the amino acids could be supplied from the medium. Moreover, the secreted proteins in the extracellular medium are degraded by the extracellular proteases, which also increases the amino acid concentration in the medium. In all the studied conditions, alanine family amino acids (alanine (Ala), leucine (Leu), valine (Val)) were not detected in the medium (Table 4.6). This is

probably due to high Ala and Leu content of EPO, which contains 13% Ala and 11% Leu, whereas the other amino acids are present at an average level of 4.2%. The other possible reason is that these are synthesized from pyruvate and pyruvate is a central and highly controlled node in the metabolism. Glutamate, isoleucine, methionine, phenyalanine and tyrosine are the other amino acids, which were never detected in the medium. Moreover, cysteine was detected in sorbitol containing media only. Aspartic acid was the amino acid secreted in highest amount in all conditions. The highest total amount of amino acid excreted to the extracellular medium did not vary according to the presence of sorbitol and reached 0.54 g L^{-1}, however varied by the methanol feed rate.

Lactic acid (Lac) was detected as the metabolic byproduct at highest amount, C_{LA}=2.6 g L^{-1} in medium without sorbitol as expected (Figure 4.42). Lactic acid is the main metabolic product formed in case of insufficient oxygen, when the TCA cycle can not take place efficiently. Since sorbitol was proposed to be used to lower the oxygen requirement of *Pichia pastoris*, the results point to the fact that this target was achieved. This was also due to shortening of the process period with sorbitol supplementation, which eliminated the lactic acid build-up period observed after 25 h of the process.

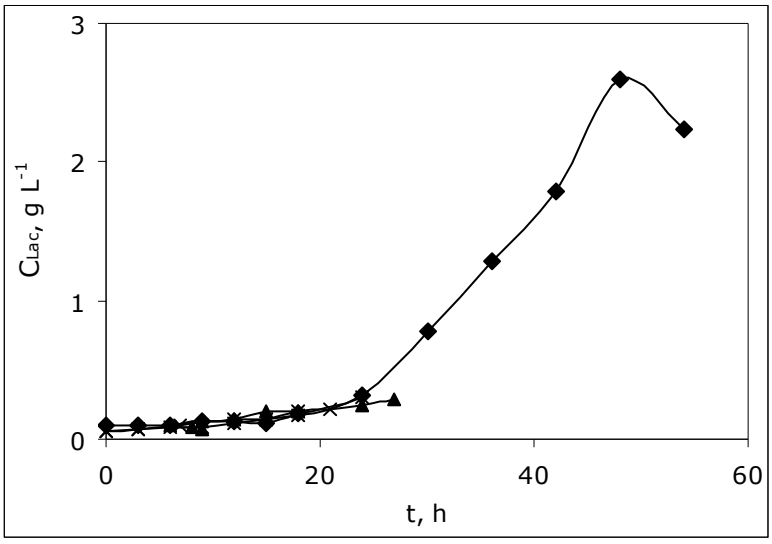


Figure 4.42 Variation in lactic acid concentration with cultivation time for different feeding profiles: M-0.03 (♦); MS-0,02 (□); MS-0,03 (▲); MS-0,04 (x).

Table 4.6 Variation in amino acid concentration* profiles with cultivation time, for runs with different feeding profiles

	RUN 1 - M-0.03 CONDITION				RUN 2 - MS-0.02 CONDITION			
Time (h)	**3**	**6**	**12**	**18**	**3**	**6**	**12**	**18**
Arg	0.036	0.000	0.034	0.000	0.018	0.000	0.002	0.000
Asn	0.072	0.021	0.051	0.040	0.040	0.038	0.029	0.034
Asp	0.118	0.022	0.072	0.048	0.071	0.045	0.073	0.037
Cys	0.000	0.000	0.000	0.000	0.000	0.000	0.033	0.000
Gln	0.075	0.016	0.050	0.046	0.040	0.030	0.040	0.041
Gly	0.033	0.005	0.027	0.011	0.020	0.012	0.013	0.013
His	0.053	0.000	0.021	0.012	0.020	0.005	0.009	0.008
Lys	0.000	0.000	0.023	0.000	0.008	0.000	0.013	0.000
Pro	0.014	0.000	0.017	0.001	0.006	0.001	0.001	0.000
Ser	0.075	0.016	0.050	0.046	0.040	0.030	0.040	0.041
Thr	0.032	0.000	0.029	0.004	0.013	0.000	0.002	0.000
Trp	0.032	0.025	0.036	0.006	0.027	0.051	0.030	0.005
Total	**0.541**	**0.106**	**0.410**	**0.215**	**0.304**	**0.212**	**0.285**	**0.179**

	RUN 3 - MS-0.03 CONDITION				RUN 4 - MS-0.04 CONDITION			
Time (h)	**3**	**6**	**12**	**18**	**3**	**6**	**12**	**18**
Arg	0.017	0.015	0.002	0.000	0.009	0.000	0.002	0.000
Asn	0.054	0.038	0.029	0.040	0.033	0.020	0.029	0.036
Asp	0.113	0.136	0.074	0.048	0.047	0.045	0.075	0.043
Cys	0.000	0.089	0.048	0.000	0.000	0.000	0.041	0.000
Gln	0.051	0.055	0.034	0.046	0.031	0.030	0.040	0.042
Gly	0.022	0.025	0.010	0.011	0.023	0.012	0.013	0.013
His	0.018	0.032	0.009	0.012	0.012	0.005	0.009	0.010
Lys	0.000	0.017	0.016	0.000	0.000	0.000	0.014	0.000
Pro	0.003	0.005	0.001	0.001	0.001	0.001	0.001	0.000
Ser	0.051	0.055	0.034	0.046	0.031	0.030	0.040	0.042
Thr	0.010	0.014	0.002	0.004	0.005	0.000	0.002	0.000
Trp	0.037	0.059	0.030	0.006	0.040	0.051	0.030	0.006
Total	**0.376**	**0.537**	**0.291**	**0.215**	**0.231**	**0.194**	**0.296**	**0.193**

* Concentrations given in g L^{-1} units

Table 4.7 Variation in organic acid concentration* profiles with cultivation time, for runs with different feeding profiles

RUN 1 - M-0.03 CONDITION													
Time (h)	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**24**	**30**	**36**	**42**	**48**	**54**
Citric acid	0.0355	0.0425	0.0828	0.1001	0.0921	0.0831	0.1149	0.1244	0.1373	0.1318	0.1578	0.2761	0.2489
Formic acid	0.0165	0.0170	0.0213	0.0722	0.1029	0.1130	0.1079	0.0634	0.0925	0.1200	0.1315	0.1465	0.1595
Fumaric acid	0.0000	0.0000	0.0006	0.0010	0.0015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Glutaric acid	0.0138	0.0160	0.0698	0.0768	0.0825	0.0837	0.0899	0.0907	0.1034	0.1028	0.1108	0.1603	0.1303
Maleic acid	0.0000	0.0000	0.0002	0.0003	0.0004	0.0004	0.0005	0.0006	0.0011	0.0012	0.0013	0.0016	0.0019
Succinic acid	0.0000	0.0000	0.0000	0.0188	0.0303	0.0340	0.0403	0.0605	0.0925	0.1262	0.1682	0.1287	0.1337

RUN 2 - MS-0.02 CONDITION						
Time (h)	**0**	**3**	**7**	**12**	**15**	**18**
Citric acid	0.0427	0.0432	0.0472	0.0437	0.0407	0.0567
Formic acid	0.0080	0.0284	0.0670	0.0791	0.0722	0.0445
Fumaric acid	0.0006	0.0004	0.0006	0.0005	0.0005	0.0000
Glutaric acid	0.0176	0.0186	0.0185	0.0201	0.0198	0.0261
Succinic acid	0.0000	0.0000	0.0486	0.0650	0.0993	0.1502

RUN 3 - MS-0.03 CONDITION									
Time (h)	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**24**	**27**
Citric acid	0.0369	0.0349	0.0393	0.0541	0.0515	0.0581	0.0693	0.0802	0.0902
Formic acid	0.0340	0.0327	0.0478	0.0611	0.0730	0.0854	0.0763	0.0449	0.0542
Fumaric acid	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0019
Glutaric acid	0.0226	0.0233	0.0252	0.0239	0.0258	0.0309	0.0328	0.0323	0.0355

RUN 4 - MS-0.04 CONDITION									
Time (h)	**0**	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**
Citric acid	0.0355	0.0425	0.0514	0.0653	0.1005	0.0944	0.1037	0.1581	0.1980
Formic acid	0.0443	0.0610	0.0690	0.0603	0.0785	0.0929	0.0274	0.0299	0.0400
Fumaric acid	0.0000	0.0000	0.0000	0.0000	0.0009	0.0000	0.0006	0.0001	0.0004
Glutaric acid	0.0195	0.0207	0.0221	0.0227	0.0325	0.0344	0.0345	0.0388	0.0547
Succinic acid	0.0000	0.0000	0.0000	0.0000	0.0215	0.0289	0.0169	0.0175	0.0513

* Concentrations given in g L^{-1} units

The other organic acids detected in the medium were mostly the TCA cycle metabolites (Table 4.7). As a general trend, more of these metabolites were secreted to the extracellular medium for the M-0.03 condition, again related to the fact that in the presence of sorbitol, the TCA cycle worked better.

Since formaldehyde is readily oxidized to formic acid or goes into the assimilatory pathway and enters glycolysis, it was never detected in the medium. Naturally, formic acid was detected in the medium in all conditions and increased with increasing methanol feed rate. On the other hand, formic acid concentration in the extracellular medium was much higher in the medium without sorbitol, meaning that formaldehyde was dissimilated rather than entering the glycolysis pathway in the absence of the co-substrate.

4.4.2.7 Oxygen Transfer Characteristics of the Bioprocess

During rHuEPO production by *P. pastoris* in fed-batch media with different methanol feeding rates, the Dynamic Method was applied to find the oxygen transfer parameters which are oxygen uptake rate (OUR), oxygen transfer rate (OTR), and oxygen transfer coefficient, K_La. At $t<0$ h, the physical oxygen transfer coefficient K_La_0 was measured in the medium in the absence of the microorganism. The variations in volumetric mass transfer coefficient (K_La), the enhancement factor E ($=K_La/K_La_0$), oxygen uptake rate (OUR), oxygen transfer rate (OTR), maximum oxygen utilization rate (OD), Damköhler number (Da) and effectiveness factor (η) throughout the bioprocesses are given in Table 4.8.

K_La first increased then decreased with the cultivation time, and the enhancement factor was always higher than 1.0 (Table 4.8). K_La depends on agitation rate, temperature, rheological properties of the fermentation medium and presence of fine particles in the mass transfer zone. Temperature and agitation rate were kept constant throughout the bioprocess, therefore the reason for the changes in K_La is due to the change in rheological properties of the fermentation medium and presence of fine particles in the mass transfer zone. As the cells grow and secrete metabolites and proteins to the extracellular medium, results in a resistance zone for mass transfer. Moreover, *Pichia pastoris* is an organism capable of high-cell density fermentation, thus cell concentration reaching 80 g L^{-1}, causing cell coalescence should have a major impact on mass transfer. On the other hand, higher cell concentration causes higher oxygen

uptake rate, which affects the volumetric mass transfer coefficient positively. Mutual effect of these two parameters (mass transfer resistance and OUR) determines the inclination of K_La throughout the fermentation process. The highest K_La values were generally obtained at t=12h of the bioprocess, which coinsides with the commence of protease secretion to the extracellular medium. Proteases are secreted for degradation of unnecessary proteins in the fermentation medium. Moreover, methanol but not sorbitol was observed to have a positive effect on K_La values.

The oxygen uptake rate (OUR), which depends on the metabolic functions and growth phases of the microorganism, tends to increase at the start of the production process, due to the increase in biomass production rate, the biomass concentration and the substrate consumption rate. OUR takes its maximum value at t=7 -12 h of the bioprocess, and thereafter decreases. Furthermore, as a general trend, presence of sorbitol lowered the OUR and increasing methanol feed rates caused an increase in the OUR. The OTR showed a similar trend to OUR but was naturally higher than OUR, since a dissolved oxygen level of DO>20% was tried to be achieved throughout the bioprocess.

The maximum possible oxygen utilization rate ($OD=\mu_{max}C_X/Y_{X/O}$) and the maximum possible mass transfer rate ($OTR_{max}=K_LaC_O^*$) were also calculated throughout the growth phase of the bioprocess. A kind of Damköhler number, Da, defined as maximum possible oxygen utilization rate per maximum mass transfer rate (Çalık et al., 2000); and effectiveness factor, η, defined as the oxygen uptake rate per maximum possible oxygen utilization rate values (Çalık et al., 2000) can be seen in Table 4.8; where η=1 is the ideal condition. It is clear in Table 4.8 that mass-transfer resistances are effective (Da>>1) at all times, rather than biochemical reaction limited (Da<1), with the highest Da= 24.1 obtained at M-0.03 condition as expected. At this condition, the cells had very high oxygen demands (OD), where as the maximum possible oxygen transfer rates (OTR_{max}) were comparatively much lower. As it is apparent from Table 4.8 that the low effectiveness factor (η) values at all times indicate that the cells are consuming lower oxygen than the oxygen demand (OD). The relatively higher η values obtained in media containing sorbitol shows that the maximum possible oxygen utilization (OD) values were better approached. Moreover, the highest η value was obtained for MS-0.03 condition, η=0.45.

Table 4.8 The variations in oxygen transfer parameters with different methanol feeding profiles

Profile Code	t (h)	K_La (s^{-1})	E K_La/K_La_0	$OTRx10^3$ ($molm^{-3}s^{-1}$)	$OTR_{max}x10^3$ ($molm^{-3}s^{-1}$)	$OURx10^3$ ($molm^{-3}s^{-1}$)	OD^x10^3 ($molm^{-3}s^{-1}$)	Da	η
M-0.03	3	0.022	2.00	16.1	23.1	14.3	-	-	-
	6	0.043	3.91	31.5	45.1	29. 7	1084.2	24.1	0.03
	12	0.067	7.27	58.7	83.8	47.3	602.2	8.6	0.08
	18	0.036	3.27	26.4	37.7	24.5	123.7	3.3	0.20
	24	0.022	2.00	16.1	23.1	14.3	76.7	3.3	0.19
	30	0.061	5.55	44.7	63.9	42.9	210.1	3.3	0.20
	36	0.032	2.91	23.5	33.5	21.6	101.5	3.0	0.21
	42	0.044	4.00	32.3	46.1	30.4	155.3	3.4	0.20
	48	0.044	4.00	32.3	46.1	30.4	153.2	3.3	0.20
MS-0.02	1	0.029	2.61	8.2	10.2	8.0	89.2	8.7	0.09
	7	0.038	3.45	10.2	12.7	10.0	28.1	2.2	0.36
	13	0.017	1.58	6.9	8.6	6.5	25.6	3.0	0.26
	15	0.018	1.63	6.3	7.9	6.0	71.8	9.1	0.08
	18	0.016	1.46	6.6	8.3	6.3	-		
MS-0.03	6	0.022	2.00	11.3	11.8	10.9	24.4	2.1	0.45
	12	0.043	3.91	22.0	27.5	21.6	65.1	2.4	0.33
	18	0.043	3.91	22.0	27.5	21.6	146.9	5.3	0.15
	24	0.014	1.27	7.2	9.0	6.8	82.6	9.2	0.08
MS-0.04	2	0.034	3.10	13.4	16.8	13.1	34.2	2.0	0.38
	9	0.045	4.09	23.4	29.2	23.0	59.3	2.0	0.39
	12	0.050	4.55	23.1	28.8	22.7	61.6	2.1	0.37
	18	0.035	3.22	18.7	23.4	18.3	105.1	4.5	0.17
	21	0.033	3.00	15.5	19.4	15.1	145.8	7.5	0.10

4.4.2.8 Yield Coefficients and Specific Rates of the Bioprocess

The overall yield coefficients give a direct idea about the efficiency and profitability of a bioprocess. The overall yield of cell generated per mass of substrate consumed ($Y_{X/S}$), the overall yield of product formed per mass of substrate consumed ($Y_{P/S}$), and the overall yield of product formed per mass of cells generated ($Y_{P/X}$) were calculated for the rHuEPO production processes with different feeding profiles (Table 4.9). For the calculation of yield coefficients related to substrate (carbon source in this study), both sorbitol and methanol were considered, since both are used as carbon sources, therefore denoted by s'.

The highest overall product yield on cell ($Y_{P/X}$=1.97 mg g^{-1}) was obtained in the medium without sorbitol (M-0.03 condition). On the other hand, using sorbitol together with low feeding rates of the toxic methanol (MS-0.02 condition), instead of using methanol only (M-0.03 condition), has increased the overall product yield on the total carbon source ($Y_{P/S'}$=1.15 mg g^{-1}) two-fold. As sorbitol and methanol have comparable prices, this also means reduction in total carbon source costs by 50%. Comparing M-0.03 and MS-0.03 conditions, 1.4-fold higher $Y_{P/S'}$ was obtained in sorbitol containing medium. Moreover, the highest cell yields on substrate was obtained at MS-0.02 condition, as $Y_{X/S'}$=0.60 g g^{-1}, which is in a consistent range with the values found in literature (Bailey and Ollis, 1986), where for most yeast and fungi grown aerobically, $Y_{X/S}$ is typically between 0.4 - 0.8 g g^{-1}, having a value of $Y_{X/S}$=0.50 g g^{-1} for *S. cerevisiae* grown aerobically on glucose. Thus, the values obtained for the cell yields on substrate are in a consistent range with the values found in literature.

The variation in total specific growth rate (μ_t), specific methanol consumption rate (q_M), specific oxygen uptake rate (q_O), specific recombinant product formation rate (q_{rp}) and specific sorbitol consumption rate (q_s) throughout the rHuEPO production processes with different feeding profiles are given in Table 4.10. Although a specific growth rate (μ_0) is aimed in fed-batch processes, it is obtained in average, but not strictly constant at the desired value. Nevertheless, for M-0.03 condition where μ_0=0.03 was aimed, the average of specific growth rates obtained throughout the process was 0.03. Thereafter, adding sorbitol batch-wise to the system without changing the feed rate, evidently resulted in higher specific growth rates, a part due to methanol (μ_0) and a part due to sorbitol (μ_S).

Table 4.9 Overall yield coefficients

Profile Code	$Y_{P/X}$ mg g^{-1}	$Y_{P/S'}$ mg g^{-1}	$Y_{X/S'}$ g g^{-1}
M-0.03	1.97	0.59	0.30
MS-0.02	1.92	1.15	0.60
MS-0.03	1.87	0.85	0.45
MS-0.04	1.88	0.73	0.39

Table 4.10 Variation in specific rates throughout the bioprocesses

Profile Code	t h	C_X g L^{-1}	μ_t h^{-1}	q_M g g^{-1} h^{-1}	q_o g g^{-1} h^{-1}	q_{rp}*1000 g g^{-1} h^{-1}	q_s g g^{-1} h^{-1}
	3	25.8	0.005	0.103	0.064	0.029	
	6	26.1	0.008	0.111	0.131	0.060	
	9	26.4	0.012	0.121		0.075	
	12	27.3	0.029	0.128	0.199	0.113	
M-0.03	18	32.4	0.034	0.129	0.087	0.101	
	24	38.4	0.040	0.128	0.043	0.088	
	30	48.0	0.039	0.123	0.103	0.080	
	36	56.4	0.037	0.118	0.044	0.079	
	42	67.5	0.040	0.113	0.052	0.070	
	48	81.0	0.027	0.113	0.043	0.043	
	54	82.5	0.017	0.137		0.036	
	3	25.8	0.078	0.077	0.038	0.148	0.104
	7	37.5	0.065	0.057	0.031	0.148	0.067
MS-0.02	12	47.3	0.039	0.050	0.016	0.107	0.068
	15	51.0	0.015	0.049	0.014	0.047	0.061
	18	50.7	0.001	0.053	0.014	0.020	0.028
	3	29.0	0.064	0.097		0.112	0.083
	6	34.5	0.066	0.089	0.017	0.152	0.078
	9	42.0	0.058	0.080		0.131	0.069
MS-0.03	12	48.0	0.054	0.077	0.051	0.134	0.073
	15	56.1	0.041	0.072		0.100	0.049
	18	60.0	0.026	0.073	0.041	0.047	0.026
	24	66.9	0.018	0.078	0.009	0.032	0.000
	27	66.9	0.007	0.084		0.029	0.000
	3	27.4	0.072	0.149	0.058	0.129	0.107
	6	35.0	0.074	0.131		0.172	0.074
	9	41.7	0.063	0.122	0.063	0.145	0.059
MS-0.04	12	49.1	0.061	0.116	0.044	0.137	0.069
	15	57.6	0.047	0.110		0.118	0.044
	18	62.5	0.028	0.113	0.034	0.060	0.020
	24	68.2	0.025	0.125	0.026	0.036	0.000

Thus, it can be assumed that the total specific growth rate is, $\mu=\mu_t=\mu_0+\mu_S$, where, $\mu_0=0.03$=constant is the target value.

The specific methanol and sorbitol consumption rates were the highest at the beginning of the process and then decreased. Both sorbitol and methanol were consumed at highest rates for MS-0.04 condition. Thus, feeding rate of methanol also affects sorbitol consumption rate, due to speeding up of the metabolism. More importantly, the highest specific recombinant protein production rate was obtained at MS-0.04 condition ($q_{rp}=0.172$ mg g^{-1} h^{-1}) (Table 4.10). It was logical to obtain increased q_{rp} with increased methanol feeding rates, since recombinant protein production is coupled to AOX enzyme in the methanol utilization pathway. Moreover, highest specific AOX activity was also obtained at t=6h of the bioprocess. On the other hand, q_{rp} towards the end of the process did not differ significantly with the feeding rate of methanol, or μ_o in this case. This was supported by the findings of Schenk et al. (2008), who reported that the specific productivity of a Mut^+ *P. pastoris* strain increases strongly with the specific growth rate for μ=0-0.02 h^{-1}, and increases only slightly with the specific growth rate above this limit. Thus, it should be kept in mind that methanol feeding rate can not be increased forever, as it is toxic, requires high amounts of oxygen supply, the maximum specific growth rate of *Pichia pastoris* on methanol is $\mu_{max}=0.14$ h^{-1}, and recalling that the forced specific growth rate of *Pichia pastoris* on methanol at MS-0.04 condition was $\mu=0.04$ h^{-1}. On the other hand, the highest specific oxygen uptake rate was obtained in the medium without sorbitol (M-0.03) as expected and moreover, the q_o values increased with increasing methanol feed rate. The methanol utilization pathway requiring oxygen is again the reason behind the above observations.

4.5 Stoichiometric Modeling of the Bioprocess and Metabolic Flux Analysis

Effects of the different feeding conditions on the intracellular reaction rate distributions were investigated by metabolic flux analysis, based on a detailed description of the stoichiometry of all relevant bioreactions involved in growth and rHuEPO formation, including the methanol and sorbitol utilization pathways, the glycolysis pathway, the gluconeogenesis pathway, the anaplerotic and the pentose phosphate pathway reactions; metabolic intermediates of the TCA cycle; the amino acid biosynthesis and catabolism pathways; biosynthesis of nucleotides, one-carbon units, carbohydrates and lipids; oxidative phosphorylation and transport reactions; and the biomass and rHuEPO synthesis reactions. The map of the metabolic network formed, consisting of m=102 metabolites and n=141 reactions is illustrated in Figure 4.43 and the stoichiometric reactions of the network used in this study are given in Appendix O. The rank of the [102x141] coefficient matrix proved that the matrix was formed from linearly independent equations.

Due to the incomplete genome project of *Pichia pastoris* and lack of metabolic studies, the intracellular biochemical reaction network of *Pichia pastoris* was constructed based on the metabolism of the model yeast *Saccharomyces cerevisiae* (Maaheimo et al., 2001; Förster et al., 2003; Kresnowati et al., 2008), the central metabolism of which has also been shown to be suitable for *Pichia stipitis* (Fiaux et al., 2003) and *Pichia pastoris* (Sola et al., 2004 and 2007). The overall pathway was simplified by lumping some reactions into single ones without loosing representation accuracy, and the details of forming the metabolic network was given in detail in Section 2.3.1.1. Also, the following major assumptions are used in the model: i) all the recombinant cells are assumed to perform identical behavior throughout the bioprocess; ii) the genetic structure, genetic control mechanism and regulation of the intracellular bioreaction network of the genetically engineered yeast are directed towards the desired protein in response to the designed environment; and, the transcription and translation steps for the protein synthesis and secretion of the protein are non-limiting; iii) the protein secretion from the recombinant *P. pastoris* is assumed to be with the passive transport mechanism; iv) the amino acid and organic acid excretion and transportation processes are assumed to proceed via the passive transport mechanism; v) CO_2, NH_3 in the form of NH_4^+, and phosphate transportations

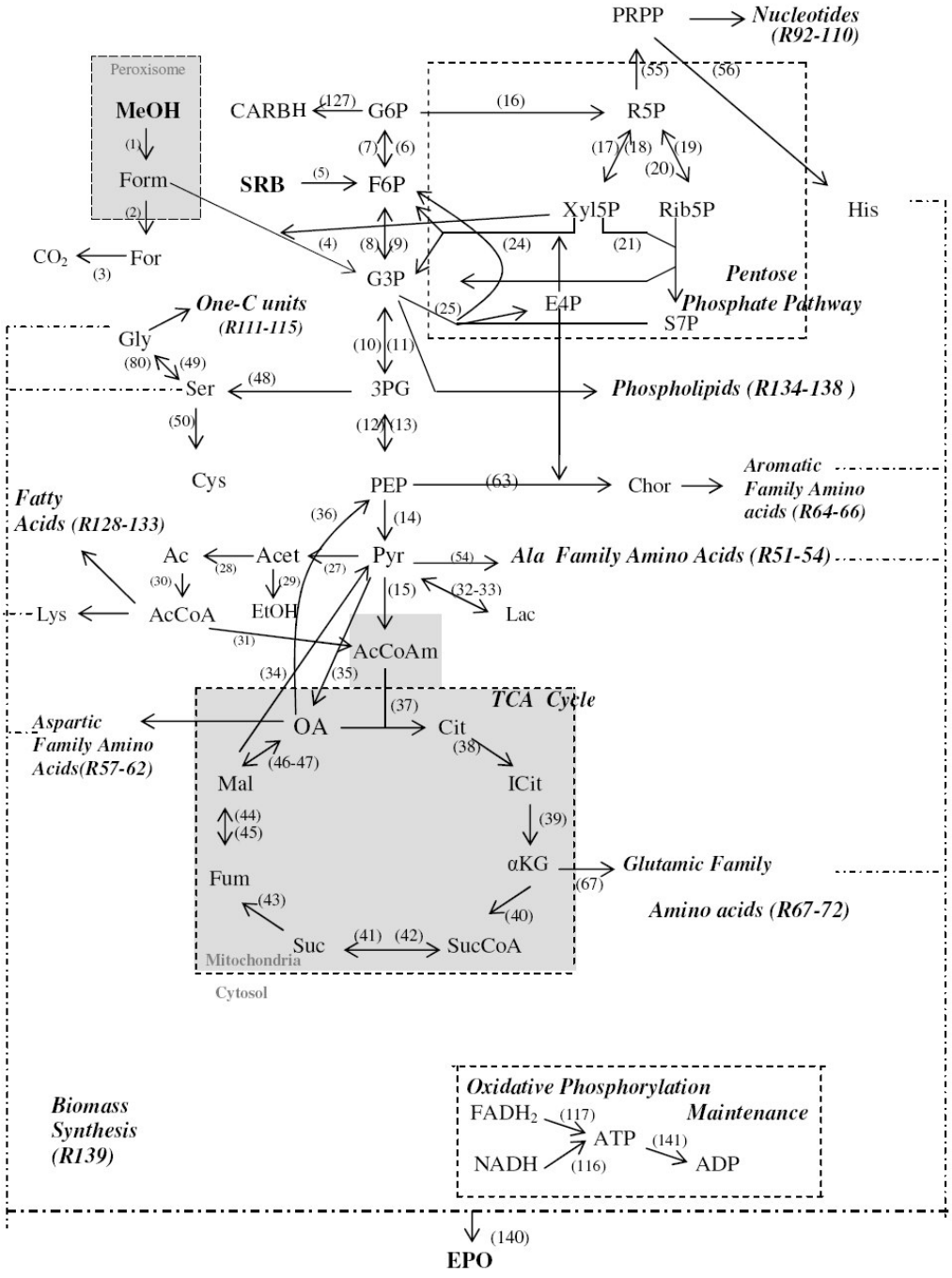


Figure 4.43 The metabolic pathway map of *P. pastoris* producing rHuEPO

from the cell to the broth and from the broth to the cell are also assumed to be with the passive transport mechanism (Çalık and Özdamar, 2002-b). vi) P/O=2 for *S. cerevisiae* (Nielsen et al., 2003) was assumed to be valid for *Pichia pastoris*; vii) compartmentation into cytosol and mitochondria was considered for metabolites (acetyl-CoA, NADPH and NADH) that cannot cross the mitochondrial membrane in both directions (Förster et al., 2002); viii) the biomass (cell) synthesis reaction (R139) was written by considering the macromolecular composition of *S. cerevisiae* (Förster et al., 2003) ix) the glyxoylate cycle was considered inactive (Sola et al., 2007; Gonzales et al., 2003). Moreover, it should be noted that, in writing EPO synthesis reaction (R140), the amino acid component and the ATP requirements for peptide bond formation between these amino acids were calculated by considering the intracellular composition of the protein, not the mature form. Thus, the prepro region, 6xHistidine and factor Xa protease recognition sequences were all taken into account. Knowing that 4 ATP is required per peptide bond formation (Mathews and van Holde, 1996), a total of 266 amino acids would require 1064 ATP (R140).

The calculated accumulation rates of the metabolites, (the organic acids, the amino acids, sorbitol, rHuEPO, and the cell) were used in the model to find the intracellular flux distributions by using equation 2.34. Considering biomass, rHuEPO, AOX and protease concentration profiles, the bioprocess can be divided into 3 periods as illustrated in Figure 4.44, for MS-0.03 condition. To analyze each period, a screenshot of the metabolism is taken, that is data from representative points of each period are used in the metabolic flux analysis. Period I ($0<t<6$ h) is the phase where AOX enzyme reaches its maximum and the rHuEPO synthesis just starts to enhance, Period II ($6<t<12$ h) is the exponential phase where the rHuEPO and cell synthesis rates were the highest, and Period III ($t>12$) is the diminution phase for rHuEPO and cell synthesis; and the start of protease production as well as the end of sorbitol consumption period. The data of cultivation times of t_1=4.5 h, t_2=9 h and t_3=15 h were used to obtain the intracellular metabolic flux distributions in Periods I, II, and III, respectively. The data obtained throughout rHuEPO production processes were exploited using a linear approach, to calculate the specific growth rate and the uptake and excretion rates of the metabolites, and these specific rates were used in the model in order to calculate the intracellular flux distributions.

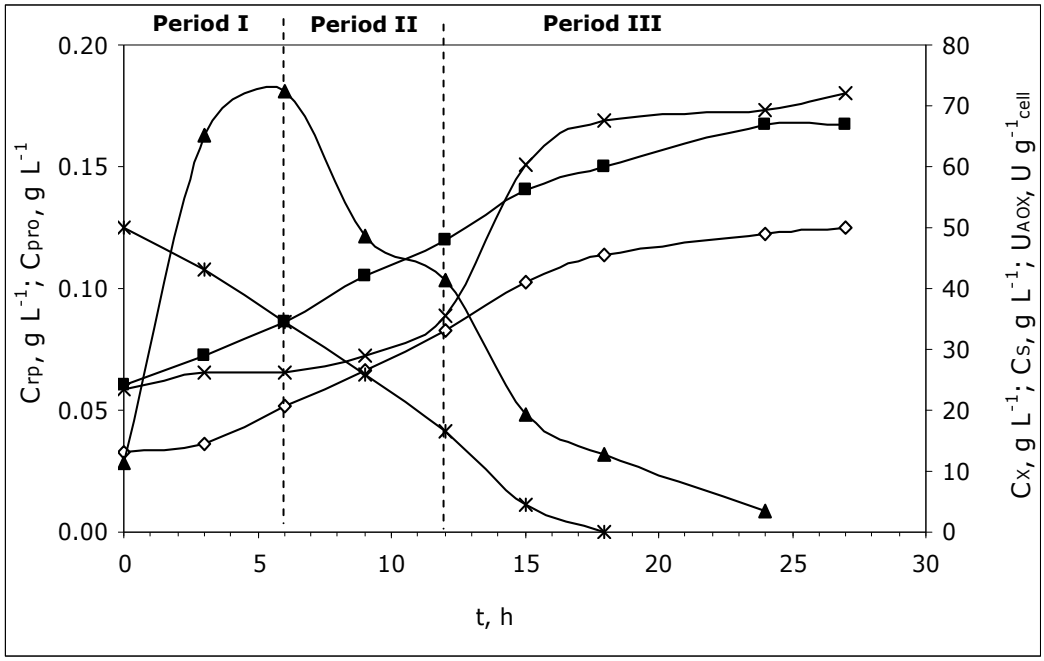


Figure 4.44 Variation in biomass and metabolite concentrations with the cultivation time for MS-0.03 condition, and the three periods of the bioprocess. RHuEPO concentration, C_{rp} (◊); protease concentration, C_{pro} (x); cell concentration, C_X (■); sorbitol concentration, C_S (*); specific AOX activity, U_{AOX} (▲).

The intracellular fluxes of the rHuEPO production processes with different feeding conditions, investigated in 3 periods, are given in Table 4.11. The first outcome from the table is that all major pathways are active at some instant, since the reactions with zero fluxes at all times (R7, 13, 15, 18, 19, 21, 23, 25, 29, 33, 42, 45, 47, 73, 78, 79, 97, 107, 119, 125) are mostly the reversible reactions, or alternative pathways. This shows that the 141 reactions represent the *Pichia pastoris* metabolism well.

From the flux results given in Table 4.11, many information can be driven such as which pathways are active and which are inactive or the increase in certain fluxes according to different conditions. Moreover, ATP generated throughout the bioprocess from different reactions is summarized in Table 4.12. Thus, each feeding condition should be analyzed in itself first, followed by comparison of the conditions at most important nodes of the metabolism, which are pyruvate (Pyr) and glyceraldehyde-3-phosphate (G3P) branch points.

Table 4.11 Metabolic flux distributions of rHuEPO production by *P. pastoris* at different feeding conditions

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
1	3.3450	3.7870	3.9980	2.0910	1.6700	1.5390	2.8850	2.5000	2.2300	4.3290	3.8260	3.4370
2		0.5660	0.4790	0.0070			0.0030	2.2660	0.0500	3.8590	0.0008	3.3300
3		0.5540	0.4790		0.0190	0.0030		2.2890	0.0670	3.8580		3.3340
4	3.3450	3.2210	3.5180	2.0840	1.6700	1.5390	2.8820	0.2340	2.1800	0.4700	3.8250	0.1070
5				0.5220	0.3280	0.2100	0.4830	0.3940	0.2690	0.6120	0.3460	0.2430
6	3.2800	1.4050	0.1500	2.0770	0.2910	0.0500	1.9600	0.3090	0.2090	0.5900	0.5820	0.1630
7												
8								0.0430				0.0700
9	3.3300	2.6280	2.4400	1.6800	0.9710	0.8460	2.1680		1.3180		2.6010	
10	0.0050	0.5730	1.0080	0.7350	0.8990	0.8810	1.0750	0.5350	1.0390	0.9180	1.2580	0.3430
11												
12	0.0050	0.0050	0.9550	0.6570	0.7590	0.8590	1.0750	0.4010	0.9420	0.7970	0.7420	0.2670
13												
14	0.0020		0.9420	0.6050	0.6930	0.8540	1.0420	0.3160	0.8910	0.7530	0.7120	0.2470
15												
16	3.2790	1.4020	0.1380	2.0370	0.2710	0.0450	1.9350	0.2870	0.1940	0.5560	0.5580	0.1470
17	3.2960	1.9980	1.2280	1.9600	0.6620	0.5330	2.1900	0.1910	0.8020	0.4470	1.4600	0.0970
18												
19												
20	0.0230	0.6100	1.1420	0.0490	0.4880	0.5020	0.3370		0.6760		1.1750	

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
21												
22	0.0230	0.6100	1.1420	0.0490	0.4880	0.5020	0.3370		0.6760		1.1750	
23												
24	0.0260	0.6130	1.1480	0.0760	0.5210	0.5050	0.3540	0.0420	0.7010	0.0220	1.1900	0.0100
25												
26	0.0230	0.6100	1.1420	0.0490	0.4880	0.5020	0.3370		0.6760		1.1750	
27						0.0170						0.0940
28						0.0170						0.0940
29												
30						0.0170						0.0940
31	0.0060	0.4810	0.0730	0.1390	0.1700	0.0330	0.8200	0.1570	0.1120	0.3350	0.4120	0.0570
32	0.0002	0.0020	0.7520	0.0030	0.3620	0.7650	0.0010	0.0130	0.6670	0.0010	0.0010	0.0020
33												
34	0.0080	0.2320	0.0770	0.0910	0.1370	0.0250	0.2980	0.1280	0.0960	0.2640	0.4540	0.0360
35		0.9870	0.1130	0.5210	0.2390	0.0490	1.5490	0.2290	0.1640	0.5460	0.4970	0.1180
36	0.0030											
37	0.0050	0.2520	0.0090	0.0960	0.0540	0.0140	0.5490	0.0540	0.0370	0.1090	0.0010	0.0400
38	0.0020	0.2520	0.0090	0.0960	0.0540	0.0140	0.5490	0.0540	0.0370	0.1080		0.0400
39	0.0020	0.2520	0.0090	0.0960	0.0540	0.0140	0.5490	0.0540	0.0370	0.1080		0.0400
40		0.0020			0.0007	0.0020	0.2440					

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
41	0.0040	0.0020	0.0005	0.0030	0.0007	0.0020	0.2440		0.0050		0.0170	
42												
43	0.0040						0.2440		0.0050		0.0160	0.0001
44	0.0140	0.2490	0.0770	0.2190	0.1370	0.0250	0.6310	0.1280	0.0960	0.2130	0.4540	0.0890
45												
46	0.0080	0.0180		0.1280			0.3330					0.0530
47										0.0500		
48		0.5680	0.0520	0.0770	0.1400	0.0220		0.1340	0.0970	0.1210	0.5170	0.0770
49	0.0050		0.0490	0.0550	0.1080	0.0200		0.0980	0.0720	0.1010	0.2670	0.0680
50	0.0000	0.0001	0.0002	0.0007	0.0010		0.0040			0.0006	0.2360	
51	0.0007	0.0040	0.0130	0.0440	0.0010	0.0050	0.0270			0.0380	0.0260	0.0180
52	0.0020	0.0040	0.0680	0.0540	0.1210	0.0210	0.0330	0.1090	0.0830	0.2100	0.4340	0.0220
53	0.0010	0.0020	0.0080	0.0260	0.0130	0.0030	0.0150	0.0140	0.0140	0.0220	0.0320	0.0100
54	0.0010	0.0020	0.0600	0.0290	0.1080	0.0180	0.0170	0.0950	0.0690	0.1880	0.4020	0.0120
55	0.0060	0.0140	0.0510	0.1270	0.0970	0.0140	0.0830	0.0950	0.0680	0.1080	0.2730	0.0510
56	0.0003	0.0010	0.0180	0.0060	0.0030	0.0007	0.0050	0.0030	0.0020	0.0050	0.0930	0.0030
57		0.7520	0.1030	0.5530	0.1850	0.0350	1.3330	0.1750	0.1270	0.3870	0.4960	0.1310
58		0.0020	0.0030	0.0100	0.0050	0.0010	0.0050	0.0050	0.0040	0.0070	0.0060	0.0040
59	0.0030	0.4950	0.0110	0.2830	0.0210	0.0050	0.9130	0.0230	0.0160	0.1300	0.0250	0.0210
60	0.0030	0.4940	0.0100	0.2780	0.0180	0.0040	0.9100	0.0200	0.0140	0.1260	0.0220	0.0190

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
61	0.0050	0.0020	0.0060	0.0220	0.0090	0.0020	0.0110	0.0100	0.0070	0.0160	0.0110	0.0080
62	0.0003	0.0005	0.0020	0.0070	0.0030	0.0008	0.0040	0.0040	0.0020	0.0060	0.0040	0.0030
63	0.0030	0.0020	0.0070	0.0260	0.0330	0.0030	0.0170	0.0420	0.0260	0.0220	0.0150	0.0100
64	0.0005	0.0010	0.0040	0.0230	0.0110	0.0030	0.0080	0.0070	0.0050	0.0110	0.0080	0.0050
65	0.0020	0.0008	0.0030				0.0060	0.0090	0.0040	0.0080	0.0060	0.0040
66	0.0009	0.0005		0.0040	0.0220		0.0030	0.0260	0.0170	0.0030	0.0010	0.0007
67		1.5890	0.2350	0.9130	0.4470	0.0900	1.7540	0.4430	0.3140	0.7750	1.1840	0.3230
68	0.0060	0.2700	0.0890	0.2690	0.1820	0.0300	0.4180	0.1760	0.1270	0.2560	0.4680	0.1090
69	3.2730	0.0020	0.0040	1.2820	0.0080	0.0020	0.0100	0.0090	0.0060	0.0130	0.0090	0.0060
70	0.0006	0.2270	0.0080	0.0300	0.0150	0.0040	0.2630	0.0170	0.0110	0.0510	0.0180	0.0120
71		0.2260	0.0040	0.0150	0.0080	0.0020	0.2540	0.0080	0.0060	0.0380	0.0090	0.0060
72	0.0010	0.0030	0.0070	0.0270	0.0180	0.0030	0.0180	0.0150	0.0100	0.0230	0.0470	0.0110
73												
74	0.0020	0.2230					0.2440			0.0260		
75	0.0040											
76	0.0007											
77						0.0004		0.0003	0.0003		0.2350	0.0002
78												
79												
80		0.2250					0.3310					

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
81			0.0170								0.0890	
82	0.0040			0.0030								
83			0.0510		0.0940	0.0140		0.0790	0.0580	0.1640	0.3850	
84				0.0100	0.0050	0.0010						
85	3.2730			1.2670								
86	0.0007	0.7700					0.3120					
87		0.4900		0.2380			0.8870			0.0950		0.0040
88	0.0010		0.0000		0.0210	0.0001		0.0250	0.0170			
89	0.0020							0.0040				
90			0.0005						0.0050		0.0170	
91					0.0040						0.0300	
92	0.0040	0.0110	0.0290	0.1020	0.0650	0.0110	0.0660	0.0580	0.0430	0.0880	0.1700	0.0410
93	0.0050	0.0130	0.0470	0.1080	0.0680	0.0120	0.0710	0.0610	0.0450	0.0930	0.2630	0.0440
94	0.0003	0.0005	0.0020	0.0060	0.0030	0.0008	0.0040	0.0040	0.0020	0.0050	0.0040	0.0030
95	0.0001	0.0001	0.0005	0.0020	0.0007	0.0002	0.0009	0.0009	0.0006	0.0010	0.0009	0.0006
96	0.0040	0.0120	0.0450	0.1020	0.0650	0.0110	0.0670	0.0570	0.0430	0.0870	0.2590	0.0410
97												
98	0.0001	0.0002	0.0007	0.0020	0.0010	0.0003	0.0010	0.0010	0.0009	0.0020	0.0010	0.0009
99	0.0001	0.0001	0.0005	0.0020	0.0007	0.0002	0.0009	0.0009	0.0006	0.0010	0.0009	0.0006
100	0.0040	0.0120	0.0450	0.1020	0.0650	0.0110	0.0670	0.0570	0.0430	0.0870	0.2590	0.0410

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
101	0.0110	0.0290	0.1010	0.2450	0.1710	0.0270	0.1630	0.1610	0.1170	0.2090	0.5430	0.0980
102	0.0006	0.0010	0.0040	0.0150	0.0070	0.0020	0.0090	0.0080	0.0060	0.0130	0.0090	0.0060
103	0.0003	0.0007	0.0030	0.0090	0.0040	0.0010	0.0050	0.0050	0.0030	0.0070	0.0050	0.0030
104	0.0003	0.0006	0.0020	0.0070	0.0030	0.0008	0.0040	0.0040	0.0030	0.0060	0.0040	0.0030
105	0.0003	0.0006	0.0020	0.0070	0.0030	0.0008	0.0040	0.0040	0.0030	0.0060	0.0040	0.0030
106	0.0002	0.0004	0.0020	0.0050	0.0030	0.0006	0.0030	0.0030	0.0020	0.0050	0.0030	0.0020
107												
108	0.0001	0.0002	0.0009	0.0030	0.0010	0.0004	0.0020	0.0020	0.0010	0.0030	0.0020	0.0010
109	0.0001	0.0002	0.0007	0.0020	0.0010	0.0003	0.0010	0.0010	0.0009	0.0020	0.0010	0.0009
110	0.0001	0.0001	0.0005	0.0020	0.0007	0.0002	0.0009	0.0009	0.0006	0.0010	0.0009	0.0006
111	0.0001	0.0001	0.0005	0.0020	0.0007	0.0002	0.0009	0.0009	0.0006	0.0010	0.0009	0.0006
112	0.0040	0.2500	0.0130	0.1640	0.0290	0.0050	0.4730	0.0250	0.0190	0.0860	0.0810	0.0200
113	0.0003	0.0005	0.0020	0.0070	0.0030	0.0008	0.0040	0.0040	0.0020	0.0060	0.0040	0.0030
114	0.0090	0.0240	0.0590	0.2100	0.1330	0.0240	0.1370	0.1190	0.0880	0.1810	0.3430	0.0850
115	0.0090	0.0240	0.0590	0.2100	0.1330	0.0240	0.1370	0.1190	0.0880	0.1810	0.3430	0.0850
116	3.3100	3.2830	1.4590	3.3360	1.4020	0.4120	4.1230	5.9630	1.1190	10.0350	3.3630	7.4900
117	3.2840		0.0520	1.2700	0.1150	0.0140	0.2440	0.1030	0.0850	0.1640	0.4180	0.0001
118	3.3140	1.7040	0.7190	1.9110	0.5150	0.0940	1.9710	2.7880	0.4180	4.6840	1.3720	3.5680
119												
120		0.3910	0.2910	1.0040	0.5740	0.1140	0.8920	0.5620	0.4000	0.8840	1.2400	0.4070

Table 4.11 Continued

	M-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.02 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.03 Condition Fluxes mmol g^{-1} DW h^{-1}			MS-0.04 Condition Fluxes mmol g^{-1} DW h^{-1}		
R #	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15	t=4.5	t=9	t=15
121	0.0140											
122	0.0002	0.0005	0.0020	0.0060	0.0030	0.0002	0.0070	0.0020	0.0010	0.0050	0.0040	0.0020
123	0.0070	0.0190	0.0760	0.1560	0.1120	0.0170	0.1040	0.1100	0.0790	0.1320	0.3810	0.0620
124	0.0130	1.0210	0.1980	0.8250	0.4320	0.0830	1.7460	0.4020	0.2900	0.8070	1.0040	0.2400
125												
126	3.3000	3.0120	1.4510	3.1120	1.3470	0.3960	2.9980	5.9080	1.0830	9.9770	3.3630	7.3980
127	0.0020	0.0030	0.0120	0.0400	0.0190	0.0050	0.0240	0.0220	0.0150	0.0340	0.0240	0.0160
128	0.0020	0.0050	0.0180	0.0610	0.0300	0.0070	0.0370	0.0340	0.0230	0.0520	0.0360	0.0250
129	0.0002	0.0005	0.0020	0.0060	0.0030	0.0007	0.0030	0.0030	0.0020	0.0050	0.0030	0.0020
130	0.0002	0.0003	0.0010	0.0040	0.0020	0.0005	0.0030	0.0020	0.0020	0.0040	0.0030	0.0020
131	0.0001	0.0002	0.0008	0.0030	0.0010	0.0003	0.0020	0.0020	0.0010	0.0020	0.0020	0.0010
132	0.0001	0.0001	0.0005	0.0020	0.0008	0.0002	0.0010	0.0009	0.0006	0.0010	0.0010	0.0007
133	0.0000	0.0000	0.0001	0.0004	0.0002	0.0000	0.0003	0.0002	0.0002	0.0004	0.0002	0.0002
134	0.0001	0.0001	0.0005	0.0020	0.0009	0.0002	0.0010	0.0010	0.0007	0.0020	0.0010	0.0007
135	0.0001	0.0001	0.0004	0.0010	0.0006	0.0001	0.0008	0.0007	0.0005	0.0010	0.0007	0.0005
136	0.0000	0.0001	0.0003	0.0010	0.0005	0.0001	0.0006	0.0006	0.0004	0.0009	0.0006	0.0004
137	0.0000	0.0000	0.0002	0.0006	0.0003	0.0001	0.0003	0.0003	0.0002	0.0005	0.0003	0.0002
138	0.0000	0.0000	0.0001	0.0005	0.0002	0.0001	0.0003	0.0003	0.0002	0.0004	0.0003	0.0002
139	0.0040	0.0080	0.0280	0.0960	0.0470	0.0110	0.0580	0.0540	0.0360	0.0810	0.0570	0.0390
140	0.0000019	0.0000035	0.0000059	0.0000087	0.0000068	0.0000020	0.0000084	0.0000062	0.0000045	0.0000103	0.0000067	0.0000052
141	3.1360			0.4690		0.2730		9.5900		16.3980		13.2440

Table 4.12 ATP generated throughout the bioprocess

	Amount of ATP Produced (mmol g^{-1} DW h^{-1})			Percent of ATP Produced (%)		
R #	**Period I**	**Period II**	**Period III**	**Period I**	**Period II**	**Period III**
M-0.03 Condition						
10	0.005	0.573	1.008	0.1	14.8	29.1
14	0.002		0.942	0.0	0.0	27.2
41	0.004	0.002	0.000	0.1	0.1	0.0
106	0.000	0.000	0.002	0.0	0.0	0.1
108	0.000	0.000	0.001	0.0	0.0	0.0
116	3.310	3.283	1.459	50.1	85.1	42.1
117	3.284		0.052	49.7	0.0	1.5
Total	**6.61**	**3.86**	**3.46**	**100**	**100**	**100**
MS-0.02 Condition						
10	0.735	0.899	0.881	12.3	28.9	40.7
14	0.605	0.693	0.854	10.2	22.3	39.5
41	0.003	0.001	0.002	0.1	0.0	0.1
106	0.005	0.003	0.001	0.1	0.1	0.0
108	0.003	0.001	0.000	0.1	0.0	0.0
116	3.336	1.402	0.412	56.0	45.0	19.0
117	1.270	0.115	0.014	21.3	3.7	0.6
Total	**5.96**	**3.11**	**2.16**	**100**	**100**	**100**
MS-0.03 Condition						
10	1.075	0.535	1.039	16.0	7.7	33.1
14	1.042	0.316	0.891	15.5	4.6	28.4
41	0.244		0.005	3.6	0.0	0.2
106	0.003	0.003	0.002	0.0	0.0	0.1
108	0.002	0.002	0.001	0.0	0.0	0.0
116	4.123	5.963	1.119	61.2	86.1	35.6
117	0.244	0.103	0.085	3.6	1.5	2.7
Total	**6.73**	**6.92**	**3.14**	**100**	**100**	**100**
MS-0.04 Condition						
10	0.918	1.258	0.343	7.7	21.8	4.2
14	0.753	0.712	0.247	6.3	12.3	3.1
41		0.017		0.0	0.3	0.0
106	0.005	0.003	0.002	0.0	0.1	0.0
108	0.003	0.002	0.001	0.0	0.0	0.0
116	10.035	3.363	7.490	84.5	58.3	92.7
117	0.164	0.418	0.000	1.4	7.2	0.0
Total	**11.88**	**5.77**	**8.08**	**100**	**100**	**100**

4.5.1 Metabolic Flux Analysis for M-0.03 Condition

The intracellular metabolic flux distribution for M-0.03 condition is given in Table 4.11. Considering the methanol metabolism, a higher fraction of formaldehyde enters the dissimilatory pathway rather than being converted to other metabolites via the assimilatory pathway. Since the entrance of the only carbon source (methanol) to the metabolism is from G3P, above this point the gluconeogenesis pathway was active, and below this point the glycolysis pathway was active. In Period II, there was a breakdown of the glycolysis and pyruvate is replenished by the anaplerotic reaction, R34. The pentose phosphate pathway (PPP) was active in all periods, however the route for the formation of the key intermediate, R5P, changed throughout the cultivation period. Moreover, its synthesis rate decreased with cultivation time.

Most of the branches from the TCA cycle were inactive. Ethanol and acetate were not detected in the medium, which was also consistent with the findings of other studies (Sola et al., 2007). While in the early periods of the bioprocess, Pyr and PEP were replenished through the anaplerotic reactions, later OA were required more. Moreover, in Period III, the TCA cycle was not completed probably due to lack of oxygen.

In Period I Ser, Asp, Asn, Glu and Arg; in Period II Gly; in Period III Trp were supplied from the medium or through catabolism reactions, since the fluxes for synthesis reactions of these amino acids were zero. In Periods II and III the flux for glutamate synthesis is the highest and for Period I, the highest flux for proline synthesis was later converted again to glutamate. Thus, the central importance of this amino acid lies in the fact that it is used for the synthesis of many other amino acids (Ser, Ala, Asp, Val, Leu, Ile, Phe, Tyr, Gln, Pro, Arg, and Lys).

Considering all the conditions investigated, only in M-0.03 condition ATP was not synthesized via oxidative phosphorylation of $FADH_2$ (Table 4.12). At this condition, ATP as mostly synthesized via oxidative phosphorylation of NADH and the rate of synthesis decreased with cultivation time. Almost half of the ATP produced was consumed for maintenance in Period I. Biomass and rHuEPO synthesis rates increased through the periods, due to the long lag phase of M-0.03 condition.

4.5.2 Metabolic Flux Analysis for MS-0.02 Condition

The intracellular metabolic flux distribution for MS-0.02 condition is given in Table 4.11. Considering the methanol metabolism, almost all of the formaldehyde enters the assimilatory pathway to be converted to other metabolites in all periods, rather than being discarded by the dissimilatory pathway. This shows that methanol feed did not exceed the cell requirements at this condition. The glycolysis, gluconeogenesis and PPP pathways worked smoothly in all periods. However, the synthesis of the key intermediate, R5P, was severely diminished throughout the cultivation.

From the anaplerotic reactions, Pyr and OA replenishment reactions were active at all times. The TCA cycle was complete only in the first period. Tyr was not synthesized throughout the bioprocess, but was supplied through other sources. Other than Tyr, all other amino acids were synthesized in Periods I and II, while Trp and Cys were also not synthesized in Period III. There was a general trend of decrease in the amino acid synthesis rates, and Glu with central importance, was again synthesized at highest amounts and with the same supply route as M-0.03 condition.

The nucleotide, carbohydrate and lipid synthesis pathway fluxes in Period I were the highest of all conditions. ATP was supplied by increasing levels of substrate-level phosphorylation reactions (R10, R14 and R41) and decreasing levels of oxidative phosphorylation reactions (R116 and R117) throughout the cultivation (Table 4.12). The lowest maintenance energy requirement of all conditions was observed at MS-0.02 condition. Moreover, the highest rHuEPO synthesis rate was obtained in Period II compared to other conditions, whereas the lowest rate was observed in the last period.

4.5.3 Metabolic Flux Analysis for MS-0.03 Condition

The intracellular metabolic flux distribution for MS-0.03 condition is given in Table 4.11. Considering the methanol metabolism, almost all of the formaldehyde enters the assimilatory pathway to be converted to other metabolites in Periods I and III, while most was discarded by the dissimilatory pathway in Period II. This fluctuation shows that the methanol feed was in excess of the metabolic requirements.

The glycolysis, gluconeogenesis and PPP pathways worked smoothly in Periods I and III, and similar to MS-0.02 condition. However, in the log phase (Period II), the glycolysis reaction from F6P to G3P was active rather than the reverse, since sorbitol enters the metabolism. This shows that the requirement for sorbitol increases in the log phase. Moreover, the PPP was mostly inactive in this period, and R5P was synthesized in low amounts. Most of the branches from the TCA cycle were inactive and from the anaplerotic reactions, Pyr and OA replenishment reactions were active at all times, similar to other MS conditions.

The TCA cycle was not completed in Periods II and III, again due to lack of oxygen. In Period I Ser, and Gly; in Periods II and III Cys and Ala were supplied from the medium or through catabolism reactions, since the fluxes for synthesis reactions of these amino acids were zero. Glutamate, the amino acid with central importance, was again synthesized at highest amounts by R67. The nucleotide synthesis pathway fluxes were the highest of all conditions in Period III.

In Period II, compared to other conditions the highest ATP synthesis rate was achieved, where most ATP was supplied by oxidative phosphorylation (Table 4.12). Moreover, in Period II, the highest maintenance energy requirements were observed, compared to other conditions. On the other hand, in the other periods, substrate-level phosphorylation had a higher impact and ATP was not wasted for maintenance energy requirements.

4.5.4 Metabolic Flux Analysis for MS-0.04 Condition

The intracellular metabolic flux distribution for MS-0.04 condition is given in Table 4.11. Considering the methanol metabolism, almost all of the formaldehyde enters the assimilatory pathway to be converted to other metabolites only in Period II, while most was discarded by the dissimilatory pathway in Periods I and III. This fluctuation shows that the methanol feed was too much at the start of the process, where the cells are still in the late adaptation phase, thereafter consumes all the methanol that is given, but then again does not require as much methanol.

The glycolysis and gluconeogenesis worked efficiently in Periods II and III, however due to preference of sorbitol consumption in Period III, the utilized route changed. PPP was mostly inactive in Periods I and III and R5P synthesis rates were relatively low in general. The TCA cycle was not completed in all the periods, thus oxygen insufficiency was the highest for this condition with the highest feeding rate of methanol, as expected.

The amino acids were most effectively synthesized at this condition, only Cys was supplied from the medium in Period III. The amino acid with central importance, Glu, was again synthesized at highest amounts by R67. The nucleotide synthesis pathway fluxes were the highest of all conditions in Period II of MS-0.04 condition. Moreover, the synthesis rates for one-carbon units, carbohydrates and lipids were the highest of all conditions in Periods II and III of MS-0.04 condition.

In Periods I and III, compared to other conditions the highest ATP synthesis rate was achieved, where most ATP was supplied by oxidative phosphorylation (Table 4.12). Moreover, at this condition the highest maintenance energy requirements were observed compared to other conditions. On the other hand, in Period II, substrate-level phosphorylation had a higher impact and ATP was not wasted for maintenance energy requirements.

4.5.5 Metabolic Flux Distribution at Specific Nodes

Analyzing the normalized fluxes at specific important nodes of the metabolism is as useful as the analysis of the fluxes given in Table 4.11, for comparison of the feeding conditions. Comparison of the normalized metabolic flux distributions at pyruvate node (in Period III), and at glyceraldehyde-3-phosphate node (in Period II), with different feeding conditions are given in Figure 4.45. The normalization was performed by the assigning 100% to the flux for the reactions R14 and R4 respectively, and calculating the other fluxes accordingly. Thus, all incoming fluxes to a node equals to all outgoing fluxes from a node, as seen in Figure 4.45. These nodes were selected for their central importance.

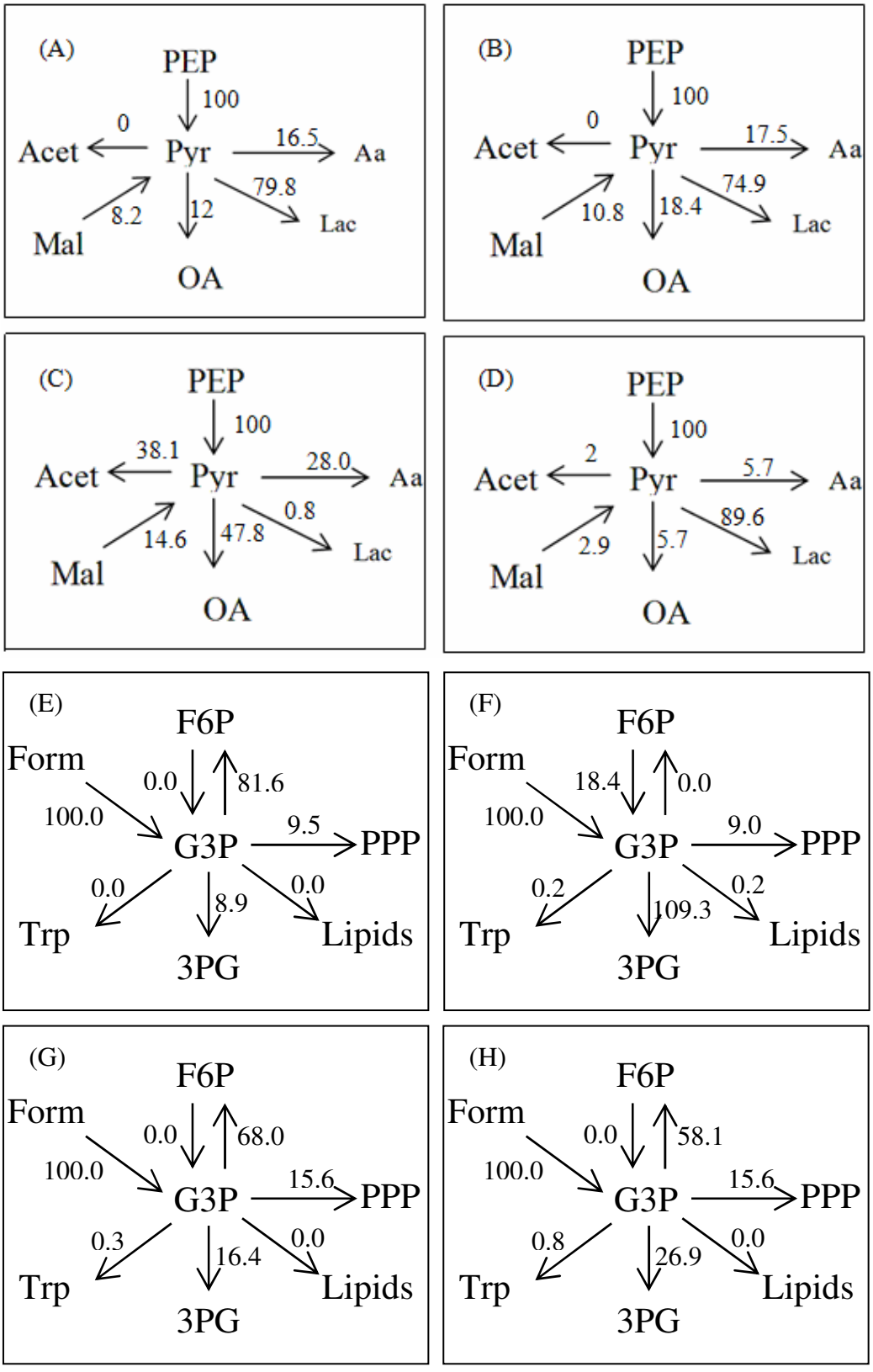


Figure 4.45 Normalized metabolic flux distributions at Pyr (at t=15 h) and G3P (at t=9 h) nodes for the bioprocesses with different with feeding conditions. A and E: M-0.03 condition; B and F: MS-0.03 condition; C and G: MS-0.04 condition; D and H: MS-0.02 condition.

The incoming flux to the pyruvate node was replenished by the anaplerotic pathway from malate, with the highest rate at MS-0.04 condition in Period III (Figure 4.45-C). At the same condition, more of the pyruvate was used in the TCA cycle and formation of certain amino acids (Ala, Val, Ile and Trp); and considering the branches of the metabolism, least was spent on byproduct formation. Moreover, the byproduct distribution was different than the other three conditions. Thus, considering only the Pyr node in Period III, MS-0.04 condition seems to be the most favorable.

Formaldehyde entering the metabolism from glyceraldehyde-3-phosphate node was supported by the other carbon source, sorbitol, at MS-0.03 condition in Period II (Figure 4.45-F), thus most were used in the glycolysis cycle, while least was used in PPP compared to the other conditions. Moreover, lipids could be synthesized in considerable amounts only at this condition.Thus, considering only the G3P node in Period II, MS-0.03 condition seems to be the most favorable.

4.6 Kinetic Modeling of the Bioprocess

4.6.1 Defining the Structure of the Model

In structured models, the first step is to define the structural or physiological origins of the model, thus the biomass is divided into a few compartments or macromolecular pools according to the physiology of the yeast and by placing cell components with similar function into the same compartment. Similar to the compartment model of Aranda et al. (2004), the cell was divided into three compartments, as shown in Figure 4.46, considering the most important effectors of recombinant protein production in *P. pastoris*. Thus, the recombinant product was considered in one compartment, alcohol oxidase and protease enzymes were gathered in the enzymatic compartment and rest of the cell components were gathered in the cellular compartment. Proteases are the enzymes that degrade proteins extracellularly or intracellularly, especially in the stationary phase of cellular growth. However, only extracellular proteases were included in the enzymatic compartment, since rHuEPO was designed to be secreted to the extracellular medium, thus it can be assumed to be affected mostly by the extracellular enzymes, and the intracellular proteases were grouped with rest of the cellular components. Moreover, it was assumed that the production rate of the extracellular proteins are equal to their transport rate to the extracellular medium, thus their intracellular accumulation rate is zero.

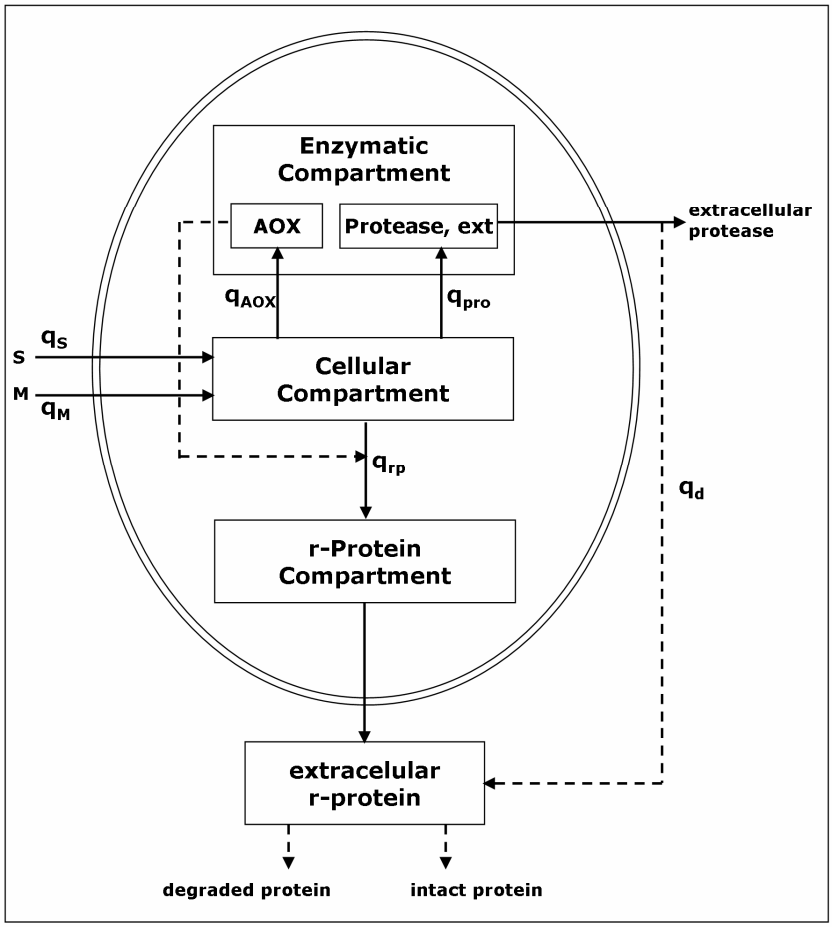


Figure 4.46 Diagram of the proposed model, showing the compartments

The fed-batch bioprocess to be modeled can be simplified as shown in Figure 4.47. AOX was not shown in extracellular medium, because it is an intracellular component. Thus, the next step is to form the material balances for each component, and then define the rate equations.

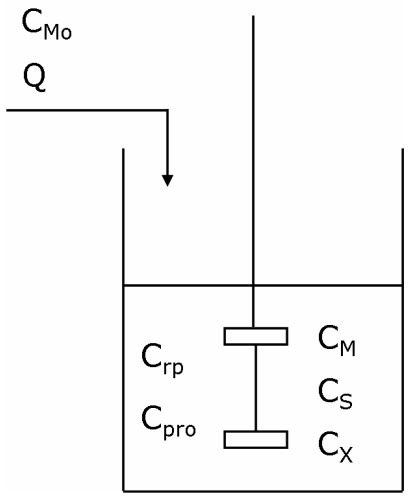


Figure 4.47 The simplified problem scheme for modeling of the fed-batch bioprocess. Q: Inlet feed rate; C_{M0}: Inlet feed concentration; C: Concentration of metabolites in the extracellular fermentation medium. Subscripts stand for: (M) methanol; (S) sorbitol; (X) cells; (rp) recombinant protein; (pro) protease.

4.6.2 Material Balances and Rate Equations

In writing the material balances, the main assumptions are that the system is isothermal and that the density of the medium is constant throughout the process, where the experimental conditions were satisfactory in achieving these criteria. Moreover, the material lost by sampling was assumed to be negligible; and that the methanol feed rate was equal to its consumption rate, thus its accumulation was assumed to be negligible. The main material balances were given in Section 2.2.

Recalling the overall mass balance, with constant density system assumption,

$$\frac{dV}{dt} = Q \tag{2.9}$$

Recalling the material balances for cell and sorbitol,

$$\frac{dC_X}{dt} = \left(\mu_t - \frac{Q}{V}\right)C_X \tag{2.10}$$

$$\frac{dC_S}{dt} = -\frac{Q}{V}C_S + q_S C_X \tag{2.14}$$

The specific growth rate in equation (2.10) was given here as the total specific growth rate (μ_t), because there are two limiting carbon sources, methanol and sorbitol.

The mass balance for the other components can be written similarly. Starting with alcohol oxidase (AOX), which is an intracellular enzyme,

$$r_{AOX}Vc = \frac{d(C_{AOX}Vc)}{dt} \tag{4.2}$$

where, V_C is the volume of the cell. Rearranging equation (4.2),

$$\frac{dC_{AOX}}{dt} = r_{AOX} - \frac{C_{AOX}}{V_C}\frac{dV_C}{dt} \tag{4.3}$$

Assuming density of the cell (ρ_c) is constant, the last term on the right in equation (4.3) can be simplified using the relation (Nielsen et al., 2003),

$$\frac{1}{V_C}\frac{dV_C}{dt} \approx \frac{1}{m_C}\frac{dm_C}{dt} \equiv \mu \tag{4.4}$$

Moreover, the substrate consumption rate (r_{AOX}) can be defined as the product of specific substrate consumption rate (q_{AOX}) and cell concentration (C_X), i.e.,

$$r_{AOX} = q_{AOX} C_X \tag{4.5}$$

Inserting equations (4.4) and (4.5) into (4.3),

$$\frac{dC_{AOX}}{dt} = q_{AOX} C_X - \mu C_{AOX} \tag{4.6}$$

In literature, the term "-μC_{AOX}" is often referred to as the intracellular dilution term and the "–QC/V" is referred to as the extracellular dilution term.

Thus, the mass balance for extracellular protease would be similar to equation (2.14) with the extracellular dilution term.

$$\frac{dC_{pro}}{dt} = q_{pro} C_X - \frac{Q}{V} C_{pro} \tag{4.7}$$

Similarly, if the degradation is neglected, the mass balance for the extracellular recombinant protein (rHuEPO) would be,

$$\frac{dC_{rp}^{*}}{dt} = q_{rp} C_X - \frac{Q}{V} C_{rp} \tag{4.8}$$

However, degradation is known to be caused by the proteases, thus should be included in equation (4.8),

$$\frac{dC_{rp}}{dt} = q_{rp}C_X - \frac{Q}{V}C_{rp} - q_d C_X \tag{4.9}$$

Thus, the material balances given in equations (2.9), (2.10), (2.14), (4.6), (4.7) and (4.9) completely define this system. However, the term simplified as Q, should be defined and related to the material balance for methanol to fix the predetermined feed rate for methanol.

Recalling the material balance for methanol given in equation (2.18), where the $C_M(dV/dt)$ term is usually negligible, the $V(dC_M/dt)$ term was also neglected considering the volume of the reactor, the production period and the low specific growth rates meaning low methanol feed rates used in this study; which all contribute to the validity of the quasi-steady state assumption (Shuler and Kargı, 2002). Moreover, methanol was never detected in the medium. Thus, considering the sensitivity of HPLC analysis, the error arising from neglecting the $V(dC_M/dt)$ term in this study was less than 1%. However, in industrial scale bioreactors, long cultivation times and high specific growth rates, neglecting this term could bring unpredictable results. Thus, after rearranging equation (2.18),

$$Q = -\frac{q_M V C_X}{C_{Mo}} \tag{4.10}$$

Assuming that methanol is mostly used for cell growth,

$$q_M = -\frac{\mu_0}{\overline{Y}_{X/M}} \tag{4.11}$$

Inserting equation (4.11) into (4.10),

$$Q = \frac{\mu_0 C_X V}{\overline{Y}_{X/M} C_{M0}} \tag{4.12}$$

It is desired to keep the methanol feeding rate proportional to exponential cell growth, and also keep specific cell growth rate constant, at $\mu_0=\mu_M=$contant, then equation (2.8) has the solution (Nielsen et al., 2003),

$$C_x V = C_{xo} V_o \exp(\mu_o t) \quad (4.13)$$

Inserting equation (4.13) into (4.12), the methanol feed rate, Q, was adjusted according to the obtained equation (Nielsen et al., 2003),

$$Q = \frac{\mu_o C_{Xo} V_o}{Y_{X/M} C_{Mo}} \exp(\mu_o t) \quad (4.14)$$

equation (4.14) can be expressed simply as,

$$Q = Q_o \exp(\mu_o t) \quad (4.15)$$

where, all the predetermined constants was denoted with Q_o,

$$Q_o = \frac{\mu_o C_{Xo} V_o}{Y_{X/M} C_{Mo}} \quad (4.16)$$

Having written the mass balances, the kinetic rate equations for the specific growth rate (μ), specific sorbitol consumption rate (q_S), specific alcohol oxidase formation rate (q_{AOX}), specific recombinant protein formation rate (q_{rp}) and specific protease formation rate (q_{pro}), have to be proposed based on yeast physiology.

Firstly, the total specific growth rate is proposed to be the sum of specific growth rate arising from methanol and sorbitol separately,

$$\mu = \mu_t = \mu_M + \mu_s \quad (4.17)$$

where, μ_M is equal to μ_o defined for equation (4.13), a desired constant specific growth rate, which is usually the design parameter of fed-batch bioprocesses.

Thus, in this study, it took values of $\mu_M=\mu_o=0.02$, 0.03 and 0.04 h^{-1}. On the other hand, μ_s was assumed to be obeying the classic Monod type kinetics,

$$\mu_s = \frac{\mu_{s,\max} C_s}{K_S + C_s} \tag{4.18}$$

Thus, total specific growth rate would be,

$$\mu_t = \mu_o + \frac{\mu_{s,\max} C_s}{K_S + C_s} \tag{4.19}$$

Sorbitol was assumed to be consumed for cell growth and maintenance, thus,

$$q_S = -\left(\frac{\mu_t}{Y_{X/S}} + m_S \right) \tag{4.20}$$

The specific AOX formation rate (q_{AOX}) should simply be related to the methanol concentration in the medium per cell concentration (C_M/C_X), since AOX is produced in response to methanol. Therefore q_{AOX} can be denoted basically as,

$$q_{AOX} = k_{a1} \frac{C_M}{C_X} + k_{a2} \tag{4.21}$$

The kinetic constant k_{a1} should be a positive value, since more methanol would mean more AOX. Magnitude of k_{a1} compared to k_{a2} shows the significance of the impact of methanol on alcohol oxidase production rate, where k_{a2} accounts for all other factors.

Protease also being in the enzymatic compartment was considered similar to the AOX enzyme, thus similar to equation (4.21),

$$q_{pro} = k_{p1} \frac{C_M}{C_X} + k_{p2} \tag{4.22}$$

The kinetic constant k_{p1} could have a negative value, because cells produce more protease in the death phase when there is a lack of nutrients. Again k_{p2} accounts for all other factors influencing protease production.

On the other hand, since r-protein production is coupled to AOX enzyme production, and thus the specific r-protein production rate should be related only to the AOX concentration per cell concentration in the medium.

$$q_{rp} = k_{rp} \frac{C_{AOX}}{C_X} \tag{4.23}$$

The kinetic constant k_{rp} shows the dependence of recombinant protein production to AOX. Thus, it should have a positive value.

The degradation of the r-protein is related to the concentration of proteases per cell concentration in the medium.

$$q_d = k_{d1} \frac{C_{pro}}{C_X} + k_{d2} \tag{4.24}$$

Again, the kinetic constant k_{d1} shows the dependence of degradation to the presence of proteases and k_{d2} accounts for all other factors.

The proposed kinetic model is complete at this point, with the equations summarized in Table 4.13 and the initial conditions given in Table 4.14. The next step is to find the kinetic constants in the model, where some can be derived from literature and some cannot, as this is the first structured model for *Pichia pastoris*.

Table 4.13 Summary of the proposed structured model equations for fed-batch production of rHuEPO by *P. pastoris*

Mass Balance Equation		**Rate Equation**	
$\frac{dC_X}{dt} = \left(\mu_t - \frac{Q}{V}\right)C_X$	(2.10)	$\mu_t = \mu_o + \frac{\mu_{s,\max} C_s}{K_S + C_s}$	(4.19)
$\frac{dC_S}{dt} = qsC_X - \frac{Q}{V}C_S$	(2.14)	$q_S = -\left(\frac{\mu_t}{Y_{X/S}} + m_S\right)$	(4.20)
$\frac{dV}{dt} = Q = Q_o \exp(\mu_o t)$ where, $Q_o = \frac{\mu_o C_{Xo} V_o}{Y_{X/M} C_{Mo}}$			(2.9, 4.15 and 4.16*)
$\frac{dC_{AOX}}{dt} = q_{AOX} C_X - \mu C_{AOX}$	(4.6)	$q_{AOX} = k_{a1}\frac{C_M}{C_X} + k_{a2}$	(4.21)
$\frac{dC_{pro}}{dt} = q_{pro} C_X - \frac{Q}{V} C_{pro}$	(4.7)	$q_{pro} = k_{p1}\frac{C_M}{C_X} + k_{p2}$	(4.22)
$\frac{dC_{rp}}{dt} = q_{rp} C_X - \frac{Q}{V} C_{rp} - q_d C_X$	(4.9)	$q_{rp} = k_{rp}\frac{C_{AOX}}{C_X}$	(4.23)
		$q_d = k_{d1}\frac{C_{pro}}{C_X} + k_{d2}$	(4.24)

* Not a rate equation.

Table 4.14 Initial conditions and fixed parameters used in the proposed structured model

Parameter	MS-0.02	MS-0.03	MS-0.04
μ_o (h^{-1})	0.02	0.03	0.04
C_{X0} (g L^{-1})	24	24	24
C_{S0} (g L^{-1})	50	50	50
C_{AOX0} (g L^{-1})	0.001	0.001	0.001
C_{rp0} (g L^{-1})	0.031	0.031	0.034
C_{pro0} (g L^{-1})	0.063	0.058	0.056
V_0 (L)	0.835	0.820	0.830

It should be noted again that, for calculating the total specific growth rate (μ_t) in the presence of sorbitol, μ_o was assumed to be constant for throughout the process at the desired value (Table 4.4 and 4.14). The fed-batch experiment performed without sorbitol (M-0.03) proved that, with the feeding profile calculated by equation (4.15) and the profile given in Figure 4.35, the desired specific growth rate (μ_o=0.03) could be achieved, when overall specific growth rates throughout the process is averaged (Table 4.10).

4.6.3 Finding the Kinetic Constants

Those kinetic constants that can be found from literature are valuable in the sense that they give the opportunity of verifying the experimental data obtained. Thus, the values for $\mu_{s,max}$, $Y_{X/S}$, $Y_{X/M}$ were experimental data reported in literature (Jungo et al., 2006 and 2007) for *P. pastoris* Mut^+ strains grown in defined medium, were assumed to be valid for the rHuEPO producing Mut^+ strain developed in this study. K_S is an intrinsic parameter of the cell-substrate system and usually has the mg per liter scale for carbohydrate substrates. Experimental

findings of Jahic et al. (2002) for Ks, determined as 0.1 g L^{-1} for both methanol and glycerol containing media was used. Since it was valid for both methanol and glycerol, it should be valid for sorbitol.

The maintenance coefficient (m_s) should be determined next. Only sorbitol was assumed to be used for maintenance energy, with the logic that it is the non-toxic substrate, and that specific growth rate does not affect maintenance energy. Maintenance coefficient was determined with one experiment, MS-0.03, as m_s=0.025 g g^{-1} h^{-1}, with the iteration results given in Figure 4.48, and it was assumed to be unaffected with different feeding conditions.

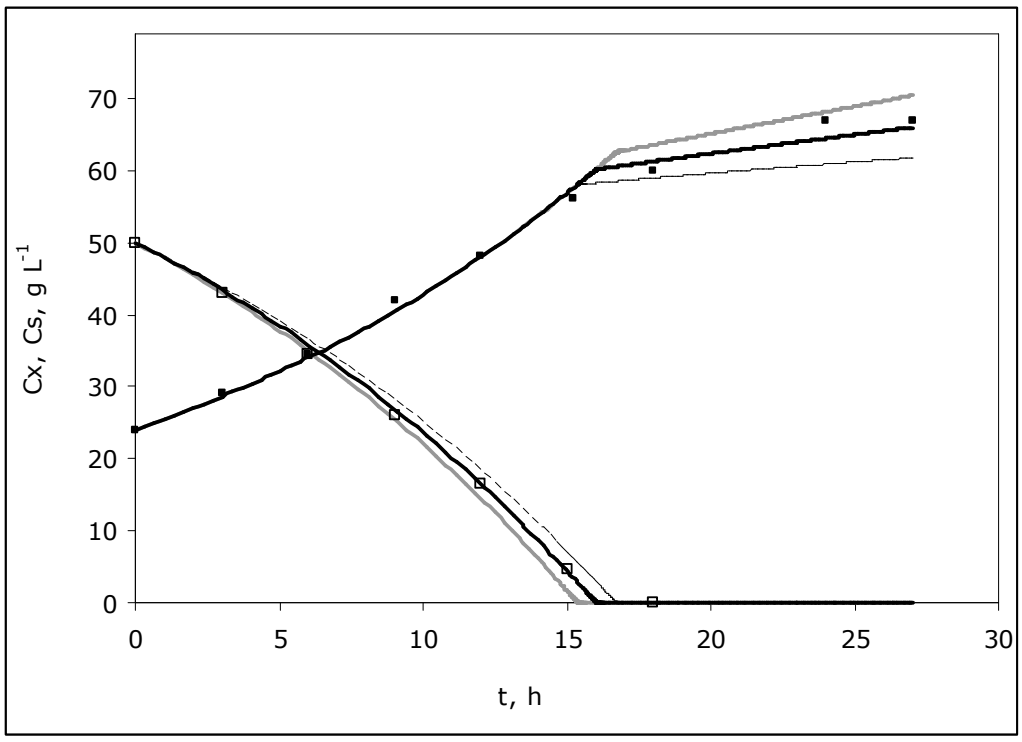


Figure 4.48. Determination of m_S by fitting the model to experimental data for MS-0.03 condition. Filled marks for C_X; empty marks for C_S. ms values: 0.02 (┈┈); 0.025 (—); 0.03 (—).

The model simulations are compared with the experimental data to estimate the kinetic constants that gives the best fit of the model to the experimental data, since the other constants can not be derived from the literature.

Starting from equation (4.21), a linear relationship between q_{AOX} and C_M/C_X should be observed to verify the equation and to be able to find the kinetic constants. Experimental value of q_{AOX} can be calculated using equation (4.6) and the instantaneous methanol concentration is known simply from the predetermined feed rate ($Q\rho_M$). A plot between q_{AOX} and C_M/C_X showed a linear relationship as expected for each feeding condition (Figure 4.49), with a coefficient of determination (R^2) larger than 0.9.

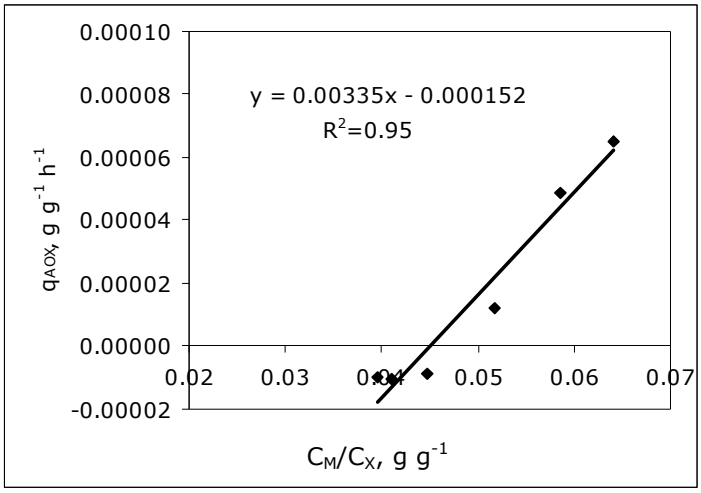
Although the kinetic constants were not the same for all conditions, a linear relationship also existed between the constants with respect to the predetermined design parameter, μ_o. The perfect linear relationship was discovered by plotting the constants k_{a1} and k_{a2} with respect to the design parameter, μ_o, set for each specific condition (Table 4.15 and Figure 4.50). Moreover, the k_{a1} values had a positive value as expected, since more methanol would mean more AOX. Magnitude of k_{a1} compared to k_{a2} shows the significance of the impact of methanol on alcohol oxidase production rate, where k_{a2} accounts for all other factors. Thus, k_{a1} was larger than k_{a2}.

In the scope of the results shown in Figures 4.49 and 4.50, the rate equation given in equation (4.21) has to be modified to reflect the relationship between the kinetic constants and the design parameter, μ_o, set for each specific condition. The new rate equation is,

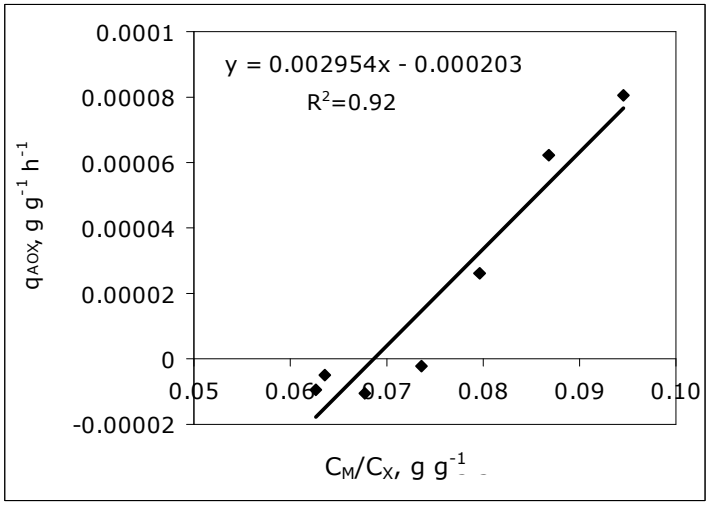
$$q_{AOX} = (k_{AOX1} - \mu_0 k_{AOX2})\frac{C_M}{C_X} - (k_{AOX3}\mu_0 + k_{AOX4}) \qquad (4.25)$$

The equations of the lines given in Figure 4.50 will give us the final kinetic constants as, k_{AOX1}=0.00419 g g^{-1} h^{-1}; k_{AOX2}=0.0425 g g^{-1} ; k_{AOX3}=0.00475 g g^{-1}; k_{AOX4}=0.058 mg g^{-1} h^{-1}.

A.



B.



C.

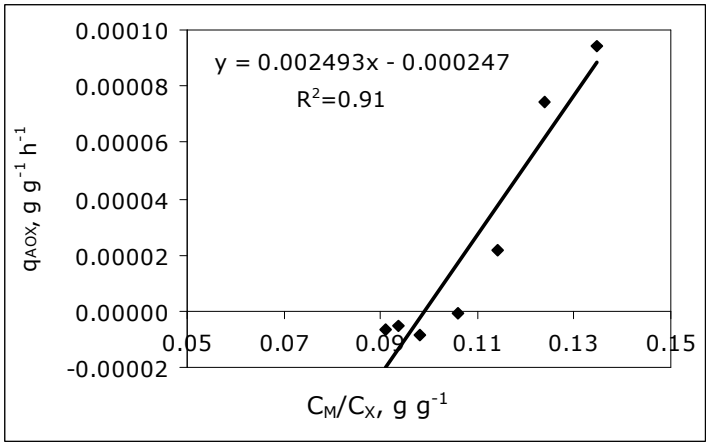
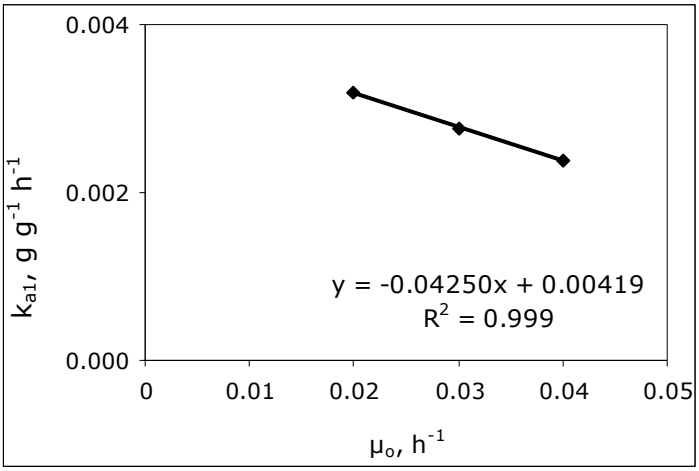


Figure 4.49 The graphs showing the rate equation correlation for specific AOX formation rate (q_{AOX}) for each fed-batch run. **A.** MS-0.02 condition **B.** MS-0.03 condition **C.** MS-0.04 condition.

Table 4.15 The kinetic constants of the specific AOX formation rate equation for different feeding conditions.

μ_o (h^{-1})	k_{a1} (g g^{-1} h^{-1})	k_{a2} (g g^{-1} h^{-1})
0.02	0.00335	0.000152
0.03	0.00290	0.000204
0.04	0.00250	0.000247

A.



B.

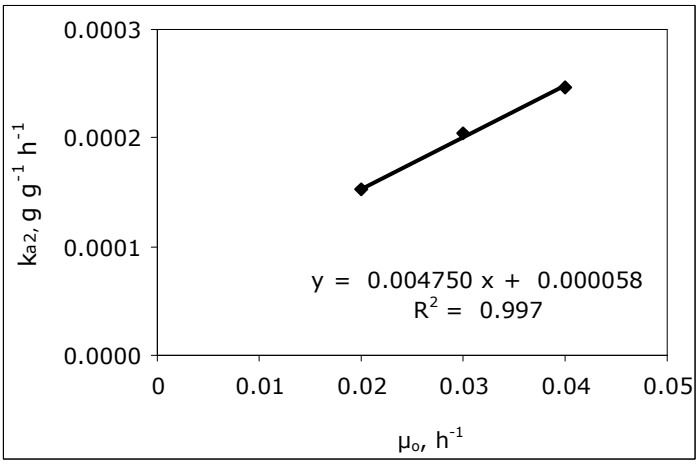


Figure 4.50 The graphs showing the linear relationship between the kinetic constants and the design parameter μ_o. **A**. The correlation graph between the kinetic constant k_{a1} and the design parameter μ_o. **B.** The correlation graph between the kinetic constant k_{a2} and the design parameter μ_o.

From a similar analysis to that performed for finding the kinetic constants in specific AOX formation rate equation, a linear relationship between q_{pro} and C_M/C_X was observed to verify equation (4.23) for each condition. As seen from Table 4.16, the same value for the first kinetic constant was obtained as k_{p1}=0.0095 g g^{-1} h^{-1}. However again the second kinetic constant was not the same for all conditions, but a linear relationship also existed between the constants with respect to the predetermined design parameter, μ_o. The perfect linear relationship was discovered again by plotting the constant k_{p2} with respect to the design parameter, μ_o, set for each specific condition (Figure 4.51).

Table 4.16 The kinetic constants of the specific protease formation rate equation for different feeding conditions.

μ_o (h^{-1})	**k_{p1}** (g g^{-1} h^{-1})	**k_{p2}** (g g^{-1} h^{-1})
0.02	-0.0095	0.00056
0.03	-0.0095	0.00080
0.04	-0.0095	0.00118

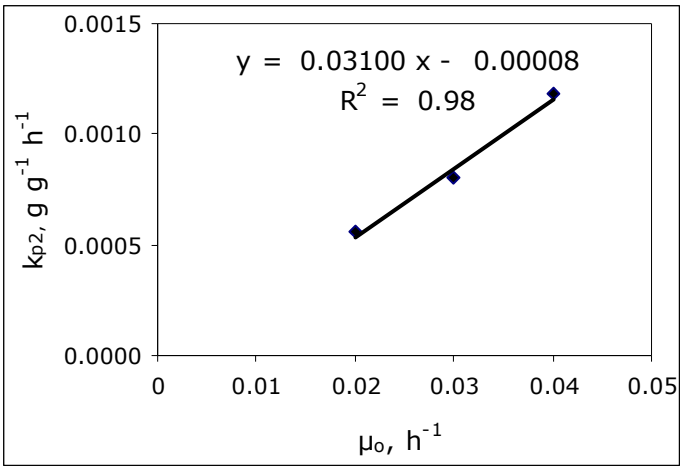


Figure 4.51 The correlation graph showing the linear relationship between the kinetic constant, k_{p2}, and the design parameter μ_o.

In the scope of the results shown in Table 4.16 and Figure 4.51, the rate equation given in equation (4.22) has to be modified to reflect the relationship between the kinetic constants and the design parameter, μ_o, set for each specific condition. The new rate equation is,

$$q_{pro} = -k_{pro1}\frac{C_M}{C_X} + (k_{pro2}\mu_0 - k_{pro3}) \qquad (4.26)$$

The equations of the line given in Figure 4.51 will give us the final kinetic constants as, $k_{p1}=k_{pro1}$ =0.0095 g g^{-1} h^{-1}; k_{pro2} =0.0310 g g^{-1}; k_{pro3} =0.00008 g g^{-1} h^{-1}. When the kinetic constants are put back into the model and the differential equation set is solved, the model represents the protease production acceptably well (Figure 4.52).

In the light of the enzymatic compartment, the recombinant protein (rHuEPO) formation rate (q_{rp}) can now be formulated. In order to find the kinetic constant in equation (4.23), the degradation of the recombinant protein has to be neglected first. The experimental q_{rp} values can be calculated from equation (4.8) and when plotted against C_{AOX}/C_X, a linear relationship is obtained (Figure 4.53).

Thus, the rate equation was formulated correctly in equation (4.23), with the kinetic constant having a value of k_{rp}=0.86 g g^{-1} h^{-1}. When the kinetic constants are put back into the model and the differential equation set is solved, the model represents the rHuEPO production with a difference (Figure 4.54), which shows the importance and requirement of including the degradation term in the model. The increase in the difference between the model and the experimental data observed after 12h of the process for MS-0.02 condition as an example in Figure 4.54, is consistent with the increase in protease production rate after 12h seen in Figure 4.52.

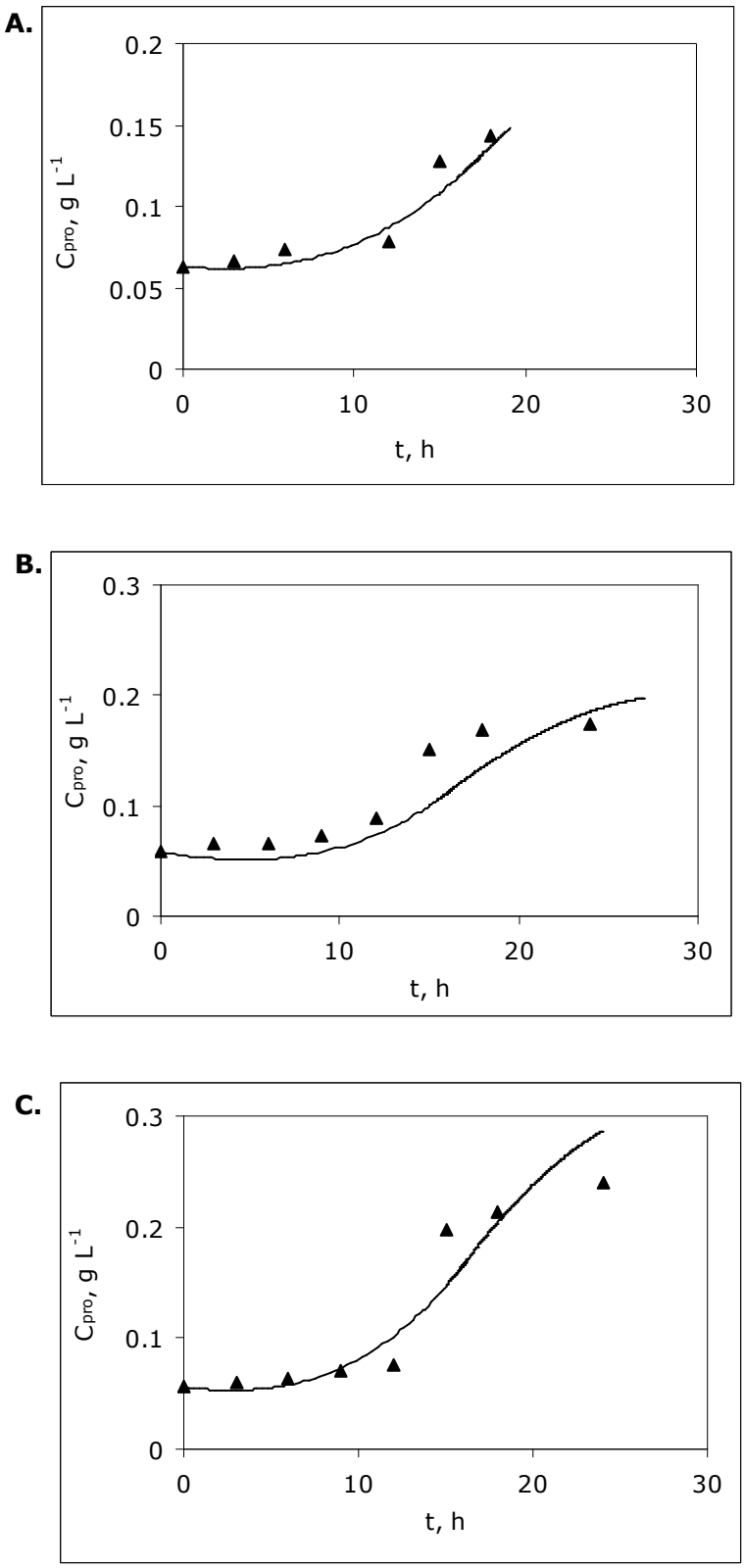


Figure 4.52 Comparison of the experimental data and modeling results of the variation in extracellular protease concentration with cultivation time. **A.** MS-0.02 condition; **B.** MS-0.03 condition; **C.** MS-0.04 condition. (▲): experimental data points; (—): model prediction.

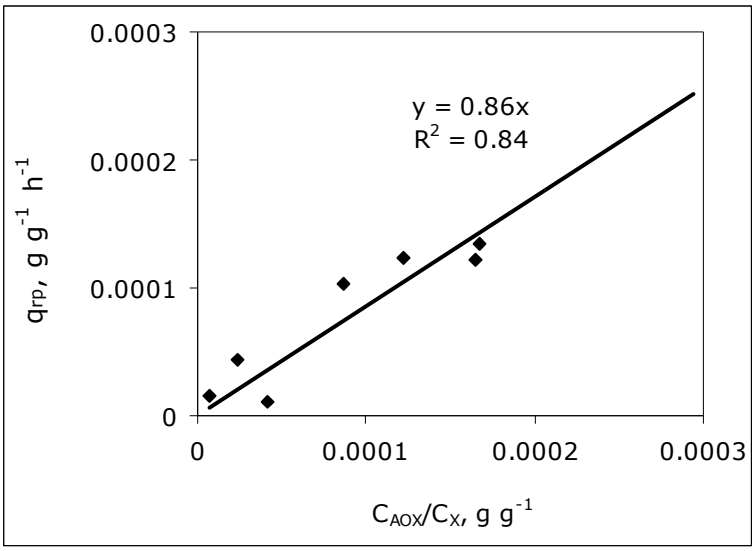


Figure 4.53 The graphs showing the rate equation correlation for specific rHuEPO formation rate (q_{rp}) for all conditions.

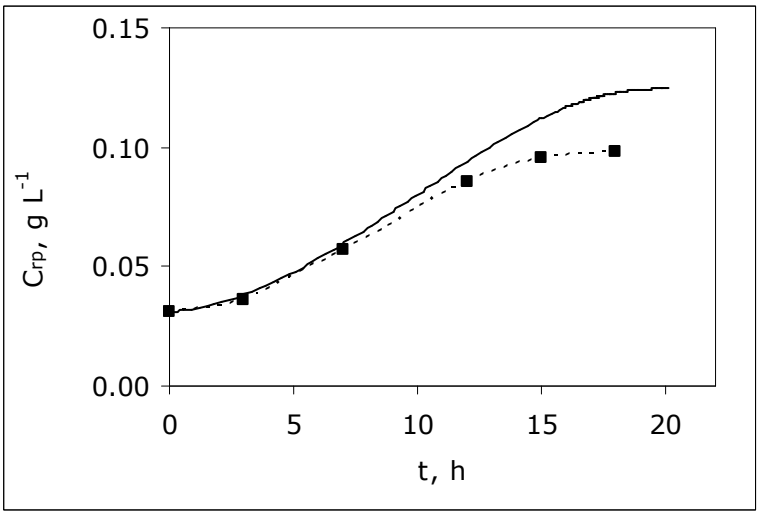


Figure 4.54 Comparison of the experimental data and modeling results of the variation in rHuEPO concentration with cultivation time for MS-0.02 condition, when degradation by proteases is neglected from the model. (■): experimental data points; (—): model prediction.

In this context, the degradation term has to be included in the model and to find the kinetic constants in equation (4.24), a linear relationship between q_d and C_{pro}/C_X should be obtained. The q_d values can be calculated from the difference between the equations (4.8) and (4.9), with the obtained k_{rp} value put into the former equation. Thus, a plot between q_d and C_{pro}/C_X showed a linear relationship as expected for each feeding condition (Figure 4.55), with a coefficient of determination (R^2) acceptable for biological modeling studies, since many simplifications have to be performed for such a complex system. Starting again from an initial analysis, demonstrated in Figure 4.55 for MS-0.04 condition, the kinetic constants were obtained for each condition after a few iterations. Although the kinetic constants were not the same for all conditions (Table 4.11), a linear relationship also existed between the constants with respect to the predetermined design parameter, μ_o. The perfect linear relationship was discovered by plotting the constants k_{d1} and k_{d2} with respect to the design parameter, μ_o, set for each specific condition (Table 4.17 and Figure 4.56). This linear relationship of the kinetic constants, thus demonstrates that the relatively low R^2 value obtained in Figure (4.55) does not show an oversimplification in the kinetic rate equation formulation, but rather that it is an initial estimate.

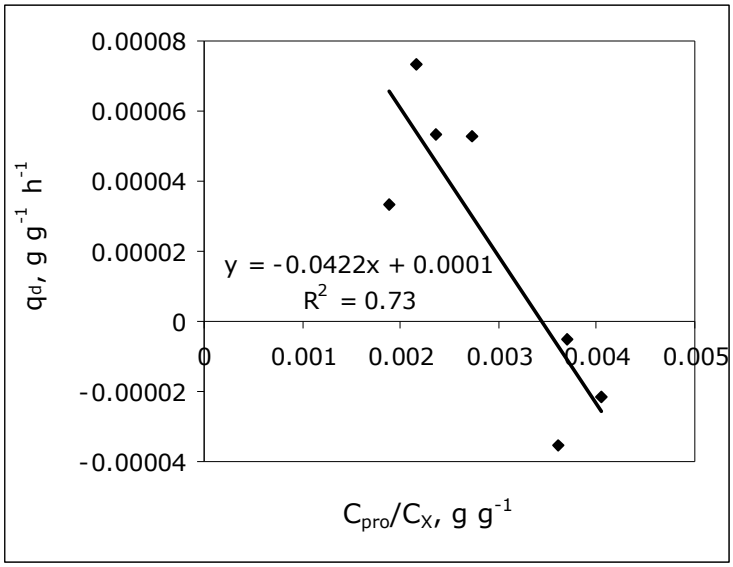


Figure 4.55 The graphs showing the rate equation correlation for specific degradation rate (q_d) for MS-0.04 condition.

Table 4.17 The kinetic constants of the specific protease formation rate equation for different feeding conditions.

μ_o (h^{-1})	k_{d1} (g g^{-1} h^{-1})	k_{d2} (g g^{-1} h^{-1})
0.02	-0.107	0.00026
0.03	-0.079	0.00017
0.04	-0.042	0.00012

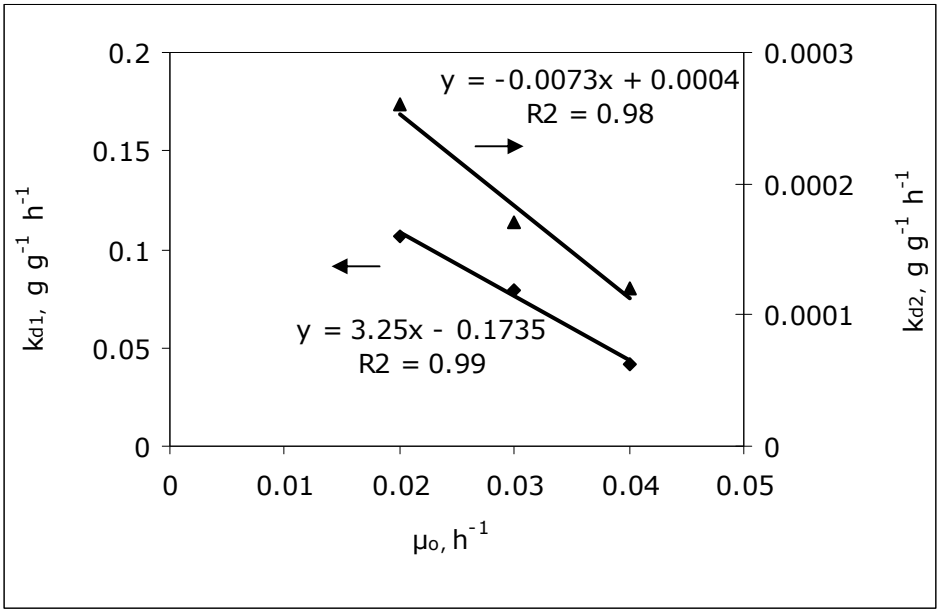


Figure 4.56 The correlation graph showing the linear relationship between the kinetic constants, k_{d1}, k_{d2} and the design parameter μ_o.

In the scope of the results shown in Table 4.11 and Figure 4.56, the rate equation given in equation (4.24) has to be modified to reflect the relationship between the kinetic constants and the design parameter, μ_o, set for each specific condition. The new rate equation is,

$$q_d = (k_{d1}\mu_0 - k_{d2})\frac{C_{pro}}{C_X} - (k_{d3}\mu_0 - k_{d4}) \qquad (4.27)$$

The equations of the lines given in Figure 4.56 will give us the final kinetic constants as, k_{d1}=3.25 g g^{-1}; k_{d2}= 0.174 g g^{-1} h^{-1}; k_{d3}=0.0073 g g^{-1}; k_{d4}=0.0004 g g^{-1} h^{-1}.

The final form of the model is summarized in Table 4.18 and the kinetic constants of the model are given in Table 4.19.

4.6.4 Validation of the Model

When the model given in Table 4.18 is solved using Polymath, with the kinetic constants given in Table 4.19 and the initial conditions given in Table 4.14, the model was found to fit the experimental data acceptably well, as seen in Figures 4.57 and 4.58. Thus, the structured kinetic model successfully predicts the cell growth, substrate consumption rate and r-product production rate in fed-batch fermentation at different design parameter (μ_0) values. Moreover, the lag period for production and the stationary phase of the production is simulated by the model, as seen in Figure 4.58, which is important to find the optimum production periods. This model can be used to predict process results of other runs with different μ_0 and C_{S0} values, which will be helpful in optimization of production without having to perform many experiments.

Table 4.18 Summary of the final structured model equations for fed-batch production of rHuEPO by *P. pastoris*

Mass Balance Equation	Rate Equation
$\frac{dC_X}{dt} = \left(\mu_t - \frac{Q}{V}\right)C_X$	$\mu_t = \mu_o + \frac{\mu_{s,\max} C_s}{K_S + C_s}$
$\frac{dC_S}{dt} = q_S C_X - \frac{Q}{V} C_S$	$q_S = -\left(\frac{\mu_s}{Y_{X/S}} + m_S\right)$
$\frac{dV}{dt} = Q = Q_o \exp(\mu_o t)$ where, $Q_o = \frac{\mu_0 C_{X0} V_0}{Y_{X/M} C_{M0}}$	
$\frac{dC_{AOX}}{dt} = q_{AOX} C_X - \mu C_{AOX}$	$q_{AOX} = (k_{AOX1} - \mu_0 k_{AOX2})\frac{C_M}{C_X} - (k_{AOX3}\mu_0 + k_{AOX4})$
$\frac{dC_{pro}}{dt} = q_{pro} C_X - \frac{Q}{V} C_{pro}$	$q_{pro} = -k_{pro1}\frac{C_M}{C_X} + (k_{pro2}\mu_0 - k_{pro3})$
$\frac{dC_{rp}}{dt} = q_{rp} C_X - \frac{Q}{V} C_{rp} - q_d C_X$	$q_{rp} = k_{rp}\frac{C_{AOX}}{C_X}$ $q_d = (k_{d1}\mu_0 - k_{d2})\frac{C_{pro}}{Cx} - (k_{d3}\mu_0 - k_{d4})$

Table 4.19 The kinetic constants of the model

Kinetic constant	**Value**
$\mu_{s,max}$	0.032 h^{-1}
Ks	0.1 g L^{-1}
m_S	0.025 g g^{-1} h^{-1}
$Y_{X/M}$	0.40 g g^{-1}
$Y_{X/S}$	0.62 g g^{-1}
k_{AOX1}	0.00419 g g^{-1} h^{-1};
k_{AOX2}	0.0425 g g^{-1}
k_{AOX3}	0.00475 g g^{-1}
k_{AOX4}	0.058 mg g^{-1} h^{-1}
k_{rp}	0.86 g g^{-1} h^{-1}
k_{pro1}	0.0095 g g^{-1} h^{-1}
k_{pro2}	0.0310 g g^{-1}
k_{pro3}	0.08 mg g^{-1} h^{-1}
k_{d1}	3.25 g g^{-1}
k_{d2}	0.174 g g^{-1} h^{-1}
k_{d3}	0.0073 g g^{-1}
k_{d4}	0.0004 g g^{-1} h^{-1}

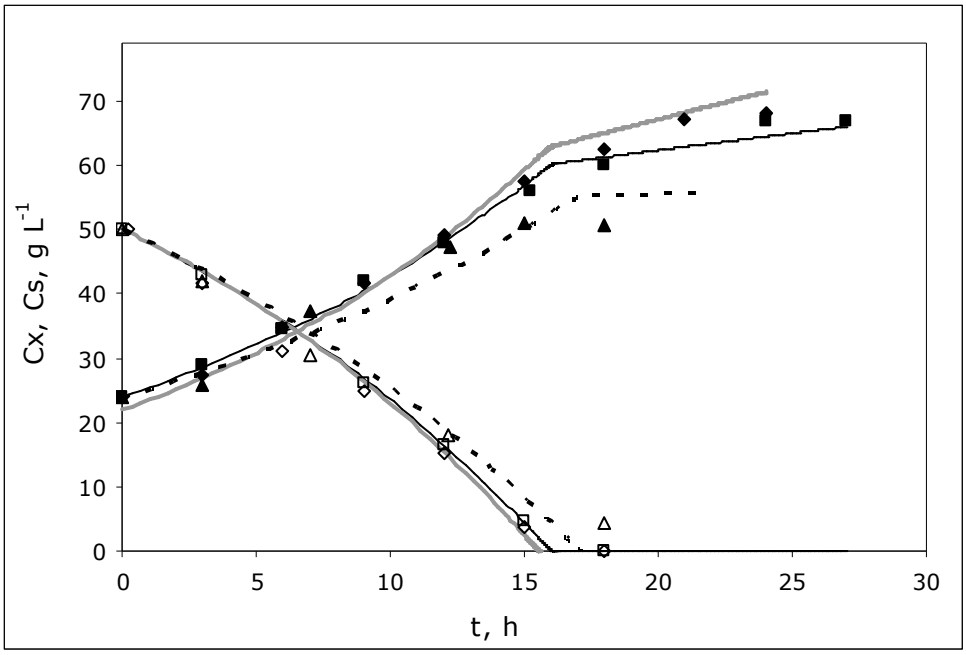


Figure 4.57 Model prediction of experimental data. Data labels: MS-0.02, (▲, Δ); MS.0.03, (■, □); MS-0.04, (♦, ◇). Filled marks for C_X; empty marks for C_S. Model: MS-0.02, (----); MS.0.03, (—); MS-0.04 (—).

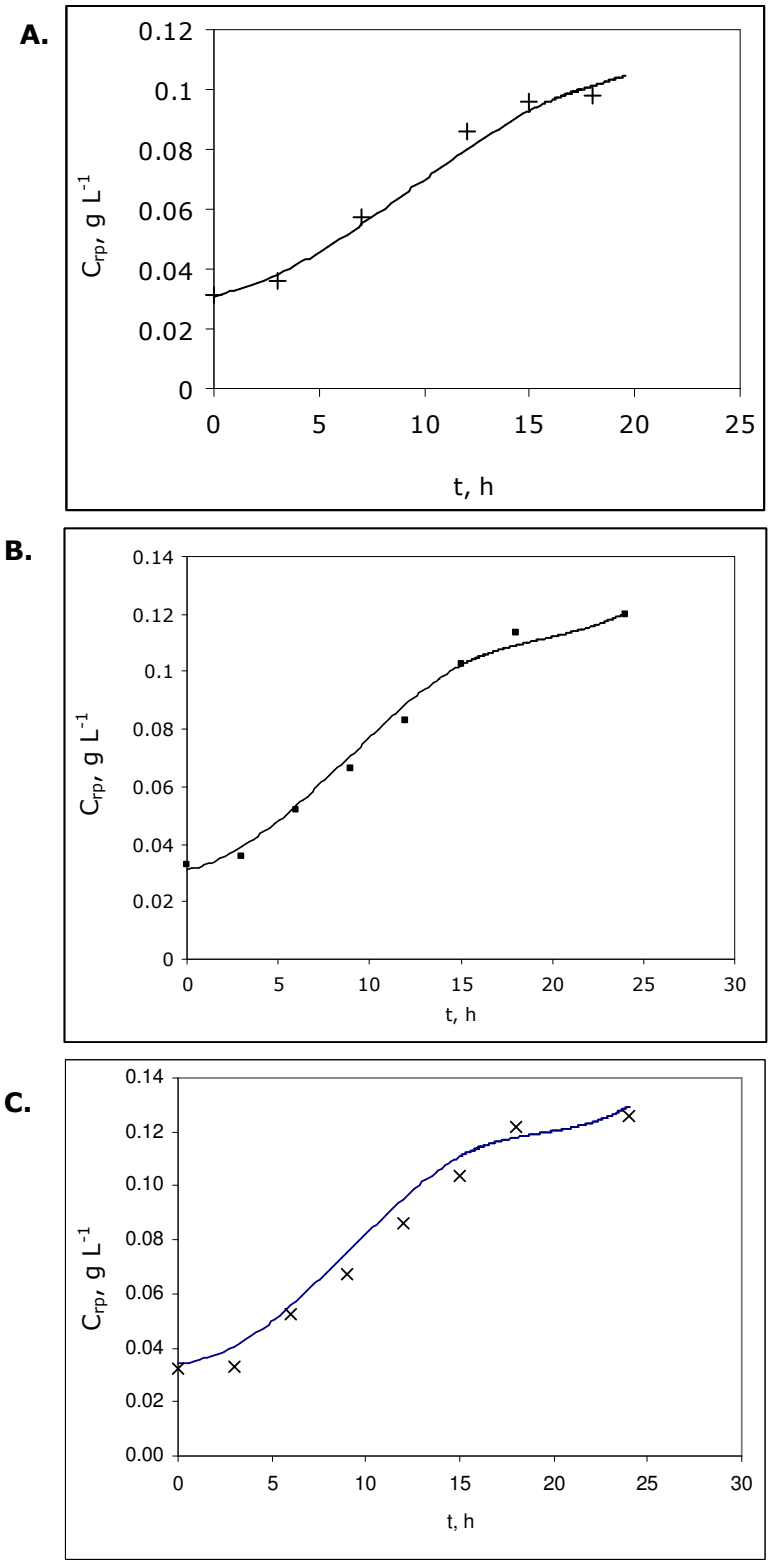


Figure 4.58 Comparison of the experimental data and modeling results of the variation in rHuEPO concentration with cultivation time. **A.** MS-0.02 condition; **B.** MS-0.03 condition; **C.** MS-0.04 condition. (x): experimental data points; (—): model prediction.

CHAPTER 5

CONCLUSION

In this study, it is aimed to design and investigate the *Pichia pastoris* expression system as an alternative to mammalian system, for the production of the therapeutically important glycoprotein, erythropoietin (EPO) and to form stoichiometric and kinetic models explaining the complex metabolic activities.

In the first step for bioprocess design, the recombinant microorganism was developed. The metabolic engineering steps within this first stage had the aims of producing the rHuEPO, that can be purified easily to a high purity and that has native *N*- and *C*-termini, which are important issues for a therapeutic protein. In this context, the coding sequence of the human EPO gene was fused with a poly-histidine affinity tag to enable rapid purification of the expressed protein product. Factor Xa protease substrate consensus sequence on the amino terminus of the wild-type *epo* sequence was also included in the cassette, to enable easy removal of all polypeptide elements that are not part of the native EPO sequence, and used as part of the final step of *in vitro* maturation. The correct construction of the pPICZαA vector was verified by restriction digestion, PCR amplification and DNA sequencing; and the integration of the constructed plasmid, pPICZαA::*epo*, into the genomic DNA of Zeocin resistant *P. pastoris* cells was verified by PCR amplification of EPO gene and Southern blotting. Thereafter, two colonies were identified as potential production systems based on their production capacities, which was detected by Western blotting. Q-ToF mass spectrometry analysis of the SDS-PAGE band of 30 kDa size from the transformant *P. pastoris* E17, verified the protein to be EPO, with 26% coverage of the amino acid sequence. More importantly, the

peptide fragments detected by MS, covered both the *N*- and *C*- termini of EPO, proving that all of the protein was being synthesized.

Once the recombinant strain for rHuEPO production was selected, the designed purification mechanism had to be verified, followed by structural verification of the purified recombinant protein. After optimization studies were carried out for affinity tag purification (i.e, on the pH of the protein solution, the amount of resin, the incubation time and temperature), 5-fold improvement in purity was achieved compared to that obtained prior to optimization studies. The purified rHuEPO, was then deglycosylated and factor Xa protease digested to obtain the native polypeptide. RHuEPO was subsequently characterized, as to the extent and nature of *N*-linked glycosylation using MALDI-ToF MS, with an analysis method that was optimized in this study. The produced glycoprotein was found to be approximately 30 kDa. All three *N*-linked glycosylation sites were occupied dominantly by $Man_{17}(GlcNAc)_2$. N-Glycanase-treated rHuEPO purified but not digested with factor Xa protease, showed a spectral peak centered about m/z 20.4 kDa, whereas the deglycosylated and factor Xa protease digested rHuEPO was 18.4 kDa, in agreement with the literature value of native EPO. Thus, the native polypeptide form of human EPO (*ca.*18kDa) was obtained for the first time in *P. pastoris* expression system, which was shown both by Western blotting and MALDI-ToF MS analyses (Çelik et al., 2007). It should be emphasized then, that the designed purification strategy was superior to the *C*-terminal polyhistidine tag already present in the pPICZαA vector, due to the fact that the factor Xa protease hydrolyzes the polypeptide on the carboxyl terminal side of the four-residue specific recognition sequence (Ile-Glu-Gly-Arg). This makes it possible to leave no residues at the *N*-terminal, following digestion, that are not part of the native sequence. Remnants of a few residues following recombinant production and maturation processes may have a significant impact on the function or stability of therapeutic proteins, including half-life. Another convenience of using factor Xa protease cleavage was that its recognition sequence (Ile-Glu-Gly-Arg) is relatively short, which makes it easy for its DNA sequence to be integrated as part of the PCR primers.

After the recombinant microorganism producing rHuEPO was developed, and the recombinant protein produced by the microorganism was purified and analyzed to verify the correct structure, the effects of medium components and pH on rHuEPO production and cell growth were investigated in laboratory scale air-

filtered shake bioreactors. Medium design studies were carried out in two distinct parts, where both complex and defined medium components were investigated.

For complex medium design, an alternative and cheap carbon source to be used in *P. pastoris* fermentation processes, the byproduct of biodiesel industry crude glycerol without prior purification, was suggested for the first time in this study (Çelik et al., 2008). Crude glycerol obtained as a byproduct from the transesterification reaction of different vegetable oils during biodiesel production, as well as different preparation techniques of the biodiesel were examined for increased efficiency of protein production by *P. pastoris*. The glycerol samples obtained from transesterification reaction using soybean and corn oils were eliminated because of their inefficiency as fermentation medium components, to avoid the white precipitate suspension formed after mixing with rest of the medium components; thus, canola and sunflower oils were used in the bioproduction medium. Cell concentration reached the maximum value in the medium containing glycerol sample C1; whereas with pure glycerol medium (G1), the lowest glycerol consumption rates were obtained which resulted in the lowest cell concentration. The results showed 1.5-fold improvement in cell concentration. Moreover, rHuEPO production by *P. pastoris* in the medium with glycerol sample C1 (from canola oil) was 1.3 times higher compared to the medium with pure glycerol and the recombinant protein concentration in the medium reached Cp = 31 g m^{-3}. Furthermore, overall cell and product yields on substrate were the highest for the glycerol sample C1, having values approximately 1.3-fold higher than pure glycerol. Canola oil is currently the most commonly used source for biodiesel production (Dmytryshyn et al., 2004), and the current study showed that, for the recombinant protein production and cell growth, canola derived glycerol byproduct of biodiesel served as the most favorable carbon source. This result was also supported by the nutritional value analysis of crude glycerol samples obtained from various vegetable oils, where canola derived crude glycerol was the most nutritious in terms of its carbon, nitrogen, fat and calorie contents (Thompson and He, 2006). Moreover, it was observed that the preparation technique of the biodiesel affected the bioprocess; and vegetable oil and methanol in a molar ratio of 1:6, and NaOH at 1% (w/v) (sample C1), were determined to be the preferable conditions from the profitability of the byproduct point of view. For this chosen condition, the biodiesel yield was estimated to be 94%; and considering the bioprocess yields,

the highest cell and product yields on substrate were obtained. Considering the product yields, the nearby alternatives (to sample C1) intended for being the favorable carbon source, S3 and C3 both had the lowest biodiesel yields; thus these crude glycerol sources should be less preferred. In this context, this study showed that, crude glycerol, the low-priced byproduct of biodiesel industry, can be efficiently used in the bioprocess by *P. pastoris* without cumbersome purification, thus demonstrating a potential use for the crude byproduct in the production of additional value-added products. If the crude glycerol obtained from fresh vegetable oils gave such promising results for the production of a therapeutic protein, process integration would be much easier for bulk chemical production.

Industrial process prefer defined medium, in cases where high purity is required in the downstream purification processes and where batch to batch variations resulting from complex media are intolerable. Thus, for the production of rHuEPO, a defined medium is more preferable, as well as for bioprocess modeling studies in general. Therefore, the effects of macro and micronutrients were investigated first, in a medium containing 10 kg m^{-3} of $(NH_4)_2SO_4$, buffered at pH=6.0, and 7.9 kg m^{-3} of methanol fed every 24 h. Rather than screening several salts individually, the most commonly used salt solutions in literature for *P. pastoris* were screened for optimizing recombinant *P. pastoris* growth. In the medium containing BSM and PTM1 salt solutions, the highest cell concentration was obtained, C_X=7.2 kg m^{-3}. Thereafter, adding sorbitol batch wise to the induction phase of *P. pastoris* was examined and in the medium containing sorbitol, methanol induction strategies were investigated. It was found that, adding sorbitol batch wise to the production medium at an initial concentration of C_{S0}= 7.5 g L^{-1}, and inducing with 1% (v/v) methanol every 24 h was found most beneficial for cellular growth and recombinant protein production in laboratory scale shake bioreactors, yielding the highest cell concentration, C_X=9.1 g L^{-1}. Since sorbitol is a non-repressing carbon source for *AOX1* promoter, sorbitol was proposed in this study to be implemented to the existing processes simply by adding the sorbitol in batch mode at the induction phase. Thus, there is no need to complicate the bioprocess by feeding the sorbitol fed-batch together with methanol. Then, the non-inhibitory concentration of sorbitol was investigated to be used in high-cell density fed-batch pilot scale bioreactor processes, where the initial sorbitol concentration optimized for laboratory scale shake bioreactors (C_{S0}=7.5 g L^{-1}) would be insufficient, and 50 g L^{-1} was found as the inhibition

limit for sorbitol. In this medium, 1.5-fold higher cell concentration was attained compared to that obtained in medium containing only methanol as the carbon source.

In the next stage of bioprocess development, where optimum bioreactor operation parameters (pH, temperature and dissolved oxygen level) had to be determined for increased cell growth and production efficiency, pH was selected as having the highest priority, since pH has a diverse effect on each specific protein, influencing proteolysis and stability of the protein. Although *P. pastoris* is capable of growing across a relatively broad pH range (3.0–7.0), pH =5.0 was found as the optimum pH for cell growth and production efficiency. A high agitation rate was chosen, N=250 rpm for laboratory scale shake-bioreactors, and N= 900 rpm for pilot scale bioreactors, since *P. pastori*s is known to require high levels of dissolved oxygen (DO). Moreover, it was observed that N=750 rpm was not sufficient at all. As *Pichia pastoris* has extremely high oxygen consumption rates and for oxygen not to be a limiting factor, usually the aim is to maintain its level above 20% saturation. Air had to be enriched with increasing amounts of oxygen, to reach these levels. Due to the mass flow controllers integrated to the oxygen and air lines entering the bioreactor system, the fluctuation in DO level was reduced about 5-fold.

The next step in the bioprocess development was to scale up to a more realistic, pilot-scale bioreactor, which is equipped with temperature, pH, foam, stirring rate, feed inlet rate and dissolved oxygen control. Four sets of fed-batch bioreactor experiments were performed, to investigate the effect of feeding rate of the carbon source, methanol, and the effectiveness of the proposed production strategy of adding sorbitol in batch-mode at the induction phase as a co-substrate. Specific growth rate (μ) being the most important design parameter in fed-batch fermentations, determines the exponential feed inlet rate of the limiting substrate methanol. The investigated feeding conditions are given in Table 4.4. Many advantages observed for the proposed production strategy, and the effects of methanol feeding rate are summarized below.

- Adding sorbitol batch wise to the medium, was beneficial in terms of eliminating the long lag-phase (9 h) for the cells and therefore cell yields were greatly improved. For instance at t=24 h, 1.7-fold higher cell concentration was obtained in medium with sorbitol (MS-0.03) compared to

that without sorbitol (M-0.03). On the other hand, "high cell density production" was achieved in as short as t=24h process, and cell concentration reached 80 g cell dry weight L^{-1} in M-0.03 condition, and 70 g CDW L^{-1} in MS-0.04 condition. These are comparable to "high cell density fermentation" claimed for *P. pastoris*, where C_X=130 g CDW L^{-1} (Wegner, 1990; Hong et al., 2002) achieved in 3-5 day processes.

- 1.8-fold higher rHuEPO was obtained at t=18 h in MS-0.04 condition compared to M-0.03 condition. Moreover, the methanol feeding rate influenced productivity as expected, since *AOX1* promoter is induced by methanol. On the other hand, rHuEPO concentration detected in the extracellular media from MS-0.03 and MS-0.04 experiments were nearly the same, yielding 130 mg L^{-1} rHuEPO.
- Methanol feeding rate did not affect sorbitol consumption rate and sorbitol utilization started immediately. Simultaneous utilization of sorbitol with methanol, confirmed the results of Jungo et al. (2007) rather than Ramon et al. (2007), who claimed sequential utilization of sorbitol and methanol.
- Regarding protease production, it was observed that as the methanol feed rate increased, the total protease concentration increased, reaching a maximum of C_{pro}=0.25 g L^{-1} in MS-0.04 condition. The obvious reason is that more cells would produce more proteases. However, when the specific protease concentrations are compared at t=24 h, C'_{pro}=2.60 mg protease g^{-1} cell dry weight for MS-0.03 condition is much less than C'_{pro}=3.06 mg protease g^{-1} cell dry weight for M-0.03 condition. Moreover, basic proteases were affected the most by the feeding rate of methanol. Thus, it can be concluded that presence of sorbitol reduces protease production.
- The specific AOX activity results showed a linear relationship to increasing feeding rates of methanol, thus at MS-0.04 condition the highest specific activity was obtained as 82 U g^{-1} CDW. The enzyme had its highest specific activity at t=6h of the bioprocess in all conditions.
- In all the studied conditions, alanine family amino acids (alanine, leucine, valine) were not detected in the medium. This is probably due to high alanine and leucine content of EPO, which contains 13% Ala and 11% Leu, whereas the other amino acids are present at an average level of 4.2%. The highest total amount of amino acid excreted to the extracellular

medium did not vary according to the presence of sorbitol and reached 0.54 g L^{-1}, however varied by the methanol feed rate.

- Lactic acid was detected as the metabolic byproduct at highest amount, C_{LA}=2.6 g L^{-1} in medium without sorbitol, as expected. Lactic acid is the main metabolic product formed in case of insufficient oxygen. Since sorbitol was proposed to be used to lower the oxygen requirement of *Pichia pastoris*, the results point to the fact that this target was achieved. This was also due to shortening of the process period with sorbitol supplementation, which eliminated the lactic acid build-up period observed after 25 h of the process.
- Oxygen uptake rate (OUR) took its maximum value at t=7-12 h of the bioprocess, and decreased thereafter. Furthermore, as a general trend, presence of sorbitol lowered the OUR and increasing methanol feed rates caused an increase in the OUR. The OTR showed a similar trend to OUR but was naturally higher than OUR. Mass-transfer resistances were effective (Da>>1) at all times, rather than being biochemical reaction limited (Da<1), with the highest Da= 24.1 obtained at M-0.03 condition as expected. At this condition, the cells had very high oxygen demands (OD), whereas the maximum possible oxygen transfer rates (OTR_{max}) were comparatively much lower. The relatively higher η values obtained in media containing sorbitol shows that the maximum possible oxygen utilization (OD) values were better approached. Moreover, the highest η value was obtained for MS-0.03 condition, η=0.45.
- Although the highest product yield on cell ($Y_{P/X}$=1.97 mg g^{-1}) was obtained in the medium without sorbitol (M-0.03 condition), using sorbitol together with low feeding rates of the toxic methanol (MS-0.02 condition), instead of using methanol only, has increased the product yield on the total carbon source ($Y_{P/S'}$=1.15 mg g^{-1}) two-fold. As sorbitol and methanol have comparable prices, this also means reduction in total carbon source costs by 50%. Moreover, the highest cell yields on substrate was obtained at MS-0.02 condition, as $Y_{X/S'}$=0.60 g g^{-1}, which is in a consistent range with the values found in literature (Bailey and Ollis, 1986).
- The highest specific recombinant protein production rate was obtained at MS-0.04 condition at t=6h (q_{rp}=0.172 g g^{-1} h^{-1}). It was logical to obtain increased q_{rp} with increased methanol feeding rates, since recombinant

protein production is coupled to AOX enzyme in the methanol utilization pathway. Moreover, highest specific AOX activity was also obtained at t=6h of the bioprocess. However, it should be kept in mind that methanol feeding rate can not be increased forever, as it is toxic, requires high amounts of oxygen supply, the maximum specific growth rate of *Pichia pastoris* on methanol is μ_{max}=0.14 h^{-1}, and recalling that the forced specific growth rate of *Pichia pastoris* on methanol at MS-0.04 condition was μ=0.04 h^{-1}.

The data collected in pilot scale bioreactor experiments were used in the next steps of bioprocess development, that is the stoichiometric and kinetic modeling steps, in order to determine the intracellular reaction rates and consequently to identify the potential metabolic bottlenecks in the synthesis of rHuEPO, as well as to better describe and predict the process outcomes. Thus, firstly the effects of the different feeding conditions on the intracellular reaction rate distributions were investigated by metabolic flux analysis, based on a detailed description of the stoichiometry of all relevant bioreactions involved. The stoichiometric model formed in this context, consisted of m=102 metabolites and n=141 reactions. The first outcome from the analysis results was that all major pathways were active at some instant, since the reactions with zero fluxes at all times were mostly the reversible reactions or alternative pathways, which showed that the 141 reactions represent the *Pichia pastoris* metabolism successfully. Considering the methanol utilization pathway, almost all of the formaldehyde enters the assimilatory pathway to be converted to other metabolites in all periods only in MS-0.02 condition, rather than being discarded by the dissimilatory pathway. This shows that methanol feed did not exceed the cell requirements at this condition. The glycolysis, gluconeogenesis and PPP pathways worked smoothly in all periods, only for MS-0.02 condition. At MS-0.02 condition, the nucleotide, carbohydrate and lipid synthesis pathway fluxes in Period I were the highest of all conditions; and the lowest maintenance energy requirements of all conditions was observed. On the other hand, not only the highest rHuEPO synthesis rate was obtained in Period II of MS-0.02, but also the lowest rate of all conditions was observed in the last period of MS-0.02 condition. In none of the feeding conditions, the TCA cycle was completed for all the periods, probably due to lack of oxygen. PPP and TCA cycles worked with highest disturbances at MS-0.04 condition, which shows the stress of increased feeding rates of methanol on cell metabolism. The amino acid with central importance, glutamate, was generally

synthesized at highest amounts and cysteine in lowest amounts. Tyrosine was not synthesized throughout the bioprocess in MS-0.02 condition, but was probably supplied through other sources. Thus, adding tyrosine to the medium in MS-0.02 condition, or cysteine in any condition, could be helpful in improving the protein synthesis pathway. Analyzing the normalized fluxes at specific important nodes of the metabolism was also found to be useful. Considering only the pyruvate node in Period III, MS-0.04 condition seemed to be the most favorable; and considering only the glyceraldehyde-3-phosphate node in Period II, MS-0.03 condition was the most favorable condition, in terms of desirable distribution of fluxes.

In the final step, the first structured kinetic model was formed for recombinant protein production with *P. pastoris*. The structural, or physiological, origins of the model was defined first, thus the biomass was divided into three compartments, or macromolecular pools, according to the physiology of the yeast and by placing cell components with similar function into the same compartment. Thus, the recombinant product was considered in one compartment, alcohol oxidase (AOX) and protease enzymes were gathered in the enzymatic compartment and rest of the cell components were gathered in the cellular compartment. Thereafter material balances were written for each component, and the kinetic rate equations were written, again based on physiology of the yeast. The kinetic constants of the model were determined based on experimental data or from literature. The developed model, made up of 6 differential equations, 6 rate equations and 17 kinetic constants, was found to fit the experimental data acceptably well. Thus, the structured kinetic model successfully predicts *P. pastoris* growth, substrate consumption rate and r-product production rate in fed-batch fermentation at different design parameter (μ_0) values. Moreover, the lag period for production and the stationary phase of the production is simulated by the model, as seen in Figure 4.58, which is important to find the optimum production periods. This model can be used to predict process results of other runs with different μ_0 or C_{S0} values, which will be helpful in optimization of production without having to perform many experiments. Moreover, the outcomes of the kinetic model can be used in MFA to further predict the intracellular reaction rates and consequently the metabolic bottlenecks.

Thus, this work contributes not only to the optimization of bioprocess design parameters for rHuEPO production by the engineered recombinant *P. pastoris*, but also to further understanding the metabolism of this methylotrophic yeast, and to developing design strategies which can be applied for the production of any bio-product using *P. pastoris*, which has become a popular microbial production system in the last decade.

REFERENCES

Agger, T. and Nielsen, J., 1999. Genetically structured modeling of protein production in filamentous fungi. Biotechnology and Bioengineering, 66 (3), 164-170.

Anumula, K.R., 2000. High-sensitivity and high-resolution methods for glycoprotein analysis. Analytical Biochemistry, 283, 17-26.

Aranda J.S., Salgado E., Taillandier P., 2004. Trehalose accumulation in *Saccharomyces cerevisiae* cells: experimental data and structured modeling. Biochemical Engineering Journal, 17, 129-140.

Araujo, A.P., Oliva, G., Henrique-Silva, F., Garrett, R.C., Caceres, O., Beltramini, L.M., 2000. Influence of the histidine tail on the structure and activity of recombinant chlorocatechol 1,2-dioxygenase. Biochemical and Biophysical Research Communications, 272, 480–484.

Azevedo A.M., Cabral J.M.S., Prazeres D.M.F., Gibson T.D., Fonseca L.P., 2004. Thermal and operational stabilities of *Hansenula polymorpha* alcohol oxidase. Journal of Molecular Catalysis B: Enzymatic, 27, 37–45.

Bailey, J.E. and Ollis, D.F., 1986. Biochemical Engineering Fundamentals. McGraw Hill Book Co., Singapore.

Bakker B.M., Overkamp K.M., van Maris A.J.A., Kötter P., Luttik, M.A.H., van Dijken J.P., Pronk J.T., 2001. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. FEMS Microbiology Reviews, 25, 15-37.

Bandyopadhyay, B. and Humprey, A.E., 1967. Dynamic measurement of the volumetric oxygen transfer coefficient in fermentation systems. Biotechnology and Bioengineering, 9, 533-544.

Bijkerk A.H.E. and Hall R.J., 1977. A mechanistic model of the aerobic growth of *Saccharomyces cerevisiae.* Biotechnology and Bioengineering, 19, 267-296.

Bill, R.M., Winter, P.C., McHale, C.M., Hoddges, Y.M., Elder, G.H., Caley, J., Flitsch, S.L., Bicknell, R., Lappin, T.R.J., 1995. Expression and mutagenesis of recombinant human and murine erythropoietins in *Escherichia coli*. Biochimica et Biophysica Acta, 1261(1), 35-43.

Blum, H., Beier, H., Gross, H.J., 1987, Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis, 8, 93-99.

Bobrowicz, P., Davidson, R.C., Li, H.J., Potgieter, T.I., Nett, J.H., Hamilton, S.R., Stadheim, T.A., Miele, R.G., Bobrowicz, B., Mitchell, T., Rausch, S., Renfer, E., Wildt, S., 2004. Engineering of an artificial glycosylation pathway blocked in core oligosaccharide assembly in the yeast *Pichia pastoris*: production of complex humanized glycoproteins with terminal galactose. Glycobiology, 14 (9): 757-766.

Bradford, M.M., 1976. A rapid and sensitive for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Analytical Biochemistry, 72, 248–254.

Brady C.P., Shimp R.L., Miles A.P., Whitmore M., Stowers A.W., 2001. High-level production and purification of P30P2MSP1$_{19}$ an important vaccine antigen for malaria, expressed in the methylotrophic yeast *Pichia pastoris*. Protein Expression and Purification, 23, 468-475.

Brierley, R.A., Bussineau, C., Kosson, R., Melton, A., Siegel, R.S., 1990. Fermentation development of recombinant *Pichia pastoris* expressing the heterologous gene: bovine lysozyme. Annals of the New York Academy of Sciences, 589: 350-362.

Brierley, R.A., Davis, R.G., Holtz, C.G., 1994. Production of insulin-like growth factor-I in methylotrophic yeast cells. US Patent No.5, 324, 639.

Bruinenberg, P.M., van Dijken, J.P., Scheffers, W.A., 1983. A theoretical analysis of NADPH production and consumption in yeasts. Journal of General Microbiology, 129, 953-964.

Burke D., Dawson D., Stearns T., 2000. Methods in yeast genetics. Cold Spring Harbor Laboratory Press, New York.

Byrne, M.P., Titball, R.W., Holley, J., Smith, L.A., 2000. Fermentation, purification, and efficacy of a recombinant vaccine candidate against botulinum neurotoxin type F from *Pichia pastoris*. Protein Expression and Purification, 18: 327–337

Callewaert, N., Laroy, W., Cadirgi, H., Geysens, S., Saelens, X., Jou, W.M., Contreras, R., 2001. Use of HDEL-tagged Trichoderma reesei mannosyl oligosaccharide 1,2-K-D-mannosidase for *N*-glycan engineering in *Pichia pastoris*. FEBS Letters 503, 173-178.

Cereghino, G.P.L. and Cregg, J.M., 1999. Applications of yeast in biotechnology: protein production and genetic analysis. Current Opinion in Biotechnology, 10, 422-427.

Cereghino, G.P.L. and Cregg, J.M., 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiology Reviews, 24, 45-66.

Cereghino, G.P.L. et al., Cereghino, J.L., Ilgen, C., Cregg, J.M., 2002. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. Current Opinion in Biotechnology, 13, 329-332.

Charoenrat, T., Ketudat-Cairns, M., Jahic, M., Veide, A., Enfors, S. O., 2005. Oxygen-limited fed-batch process: an alternative control for *Pichia pastoris* recombinant protein processes, Bioprocess and Biosystems Engineering, 27: 399–406.

Chauhan, A.K., Arora, D., Khanna, N., 1999. A novel feeding strategy for enhanced protein production by fed-batch fermentation in recombinant *Pichia pastoris.* Process Biochemistry, 34:139–145.

Chen, Y., Jin, M., Egborge, T., Coppola, G., Andre, J., Calhoun, D.H., 2000. Expression and characterization of glycosylated and catalytically active recombinant human a-galactosidase A produced in *Pichia pastoris*. Protein Expression and Purification, 20: 472-484.

Chen, Y.H., Cai, R.X., Zhang, K., 2007. Study the Effect of Vitamin Kon Intracellular NAD Level in Yeast by Fluorescence Spectrum. Spectrochimica Acta, Part A, 67, 235-239.

Chiruvolu, V., Cregg, J.M. and Meagher, M.M., 1997. Recombinant protein production in an alcohol oxidase-defective strain of *Pichia pastoris* in fedbatch fermentations. Enzyme and Microbial Technology, 21(4): 277-283.

Choi, B.K., Bobrowicz, P., Davidson, R.C., Hamilton, S.R., Kung, D.H., Huijuan, L., Miele, R.G., Nett, J.H., Wildt, S., Gerngross, T.U. (2003). Use of combinatorial genetic libraries to humanize *N*-linked glycosylation in the yeast *Pichia pastoris*. Proceedings of the National Academy of Sciences USA, 100(9), 5022-5027.

Chung, J.D., 2000. Design of metabolic feed controllers: Application to high density fermentation of *Pichia pastoris:* Mathematical model for fedbatch high cell density culture. Biotechnology and Bioengineering, 68:298–307.

Cindric, I.J., Zeiner, M., Steffan, I., 2007. Trace Elemental Characterization of Edible Oils by ICP-AES and GFAAS. Microchemical Journal, 85, 136-139.

Clare, J.J., Rayment, F.B., Ballantyne, S.P., Sreerkrishna, K., Romanos, M.A., 1991. High-level expression of tetanus toxin fragment C in *Pichia pastoris* strains containing multiple tandem integrations of the gene. BioTechnology, 9: 455–460.

Cohen, A., 1983. The PICO TAG System: a New Method to Analyze Primary and Secondary Amino Acids with One Picomole Sensitivity. Biotechniques, Sept/Oct, 273-275.

Conor McCarthy, Technelysium Pty Ltd, http://www.technelysium.com. au/chromas.html, Last accessed: June 2008.

Cortassa S., Aon J.C., Aon M.A., 1995. Fluxes of carbon, phosphorylation and redox intermediates during growth of *Saccharomyces cerevisiae* on different carbon sources. Biotechnology and Bioengineering, 47, 193- 208.

Cos O., Resina D., Ferrer P., Montesinos J.L., Valero F., 2005. Heterologous production of Rhizopus oryzae lipase in *Pichia pastoris* using the alcohol oxidase and formaldehyde dehydrogenase promoters in batch and fed-batch cultures. Biochemical Engineering Journal, 26, 86-94.

Cos, O., Ramon, R., Montesinos, J.L., Valero, F., 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. Microbial Cell Factories, 5(1), 17.

Couderc, R. and Baratti, J. (1980). Oxidation of methanol by the yeast, *Pichia pastoris*. Purification and properties of the alcohol oxidase. Agricultural Biology and Chemistry, 44(10), 2279-2289.

Cregg, J.M., Madden, K.R., Barringer, K.J., Thill, G.P., Stillman, C.A., 1989. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. Molecular and Cellular Biology 9(3), 1316-1323.

Cregg, J.M., Vedvick, T.S., Raschke, W.C., 1993. Recent Advances in the Expression of Foreign Genes in *Pichia pastoris*. BioTechnology, 11 (8), 905-910.

Cregg, J.M., 1999. Gene expression systems. Edited by Fernandez, J.M. and Hoeffler, J.P., Academic Press, Sen Diego, California.

Cunha A.E., Clemente J.J., Gomes R., Pinto F., Thomaz M., Miranda S., Pinto F., Moosmayer D., Donner P., Carrondo M.J.T., 2004. Methanol induction optimization for scFv antibody fragment production in *Pichia pastoris*. Biotechnology and Bioengineering, 86 (4), 458–467.

Curvers, S., Brixius, P., Klauser, T., Thommes, J., Weuster-Botz, D., Takors, R., Wandrey, C., 2001-a. Human chymotrypsinogen B production with *Pichia pastoris* by integrated development of fermentation and downstream processing. Part 1. Fermentation. Biotechnology Progress, 17:495–502.

Curvers, S., Linnemann, J., Klauser, T., Wandrey, C. and Takors, R., 2001-b. Recombinant protein production with *Pichia pastoris* in continuous fermentation - Kinetic analysis of growth and product formation. Chemie Ingeniur Technik, 73: 1615-1621.

Çalık, P., 1998. Bioprocess development for serine alkaline protease production. PhD thesis, Ankara University, Ankara.

Çalık, P., Çalık, G., Takaç, S., Özdamar, T.H., 1999. Metabolic Flux Analysis for Serine Alkaline Protease Fermentation by *Bacillus licheniformis* in a Defined

Medium: Effects of the Oxygen Transfer Rate, Biotechnology and Bioengineering, 64, 151-167.

Çalık, P., Çalık, G., Özdamar, T.H., 2000. Oxygen Transfer Strategy and its Regulation Effects in Serine Alkaline Protease Production by *Bacillus licheniformis*, Biotechnology and Bioengineering, 69, 301-311.

Çalık, P. and Özdamar, T.H., 2002-a. Bioreaction network flux analysis for industrial microorganisms: A review. Reviews in Chemical Engineering, 18 (6), 553-596

Çalık P. and Özdamar T.H., 2002-b. Metabolic flux analysis for human therapeutic protein productions and hypothesis for new therapeutical strategies in medicine. Biochemical Engineering Journal, 11 (1), 49-68.

Çalık, P., Tomlin, G., Oliver, S.G., Özdamar, T.H., 2003. Overexpression of Serine Alkaline Protease in Bacillus licheniformis and its Impact on the Metabolic Reaction Network, Enzyme and Microbial Technology, 32(6), 706-720.

Çelik, E., Çalık, P., Halloran, S.M., Oliver, S.G., 2007. Production of recombinant human erythropoietin from *Pichia pastoris* and its structural analysis, Journal of Applied Microbiology, 103(6), 2084-2094.

Çelik, E., Özbay, N., Oktar, N., Çalık, P, 2008. Use of biodiesel byproduct crude glycerol as the carbon source for fermentation processes by recombinant *Pichia pastoris.* Industrial & Engineering Chemistry Research, 47, 2985-2990.

Daly, R. and Hearn, M.T.W., 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. Journal of Molecular Recognition, 18, 119–138.

de Almeida, J.R.M., de Moraes, L.M.P., Torres, F.A.G., 2005. Molecular characterization of the 3-phosphoglycerate kinase gene (PGK1) form the methylotrophic yeast *Pichia pastoris*. Yeast 22, 725-737.

Dell, A. and Morris, H.R., 2001. Glycoprotein Structure Determination by Mass Spectrometry. Science 291, 2351-2356.

Doran, P.M., 1995. Bioprocess Engineering Principles, Academic Press, San Diego.

Dmytryshyn, S.L., Dalai, A.K., Chaudhari, S.T., Mishra, H.K., Reaney, M., 2004. Synthesis and Characterization of Vegetable Oil Derived Esters: Evaluation for their Diesel Additive Properties. Bioresource Technology, 92, 55-64.

Egli, T., van Dijken, J.P., Veenhuis, M., Harder, W. and Fiechter, A., 1980). Methanol metabolism in yeasts : regulation of the synthesis of catabolic enzymes. Archives of Microbiology, 124: 115-121.

Egli, T., Käppeli, O. and Fiechter, A. 1982-a. Mixed substrate growth of methyltrophic yeasts in chemostat culture : influence of the dilution rate on the utilisation of a mixture of glucose and methanol. Archives of Microbiology, 131: 8-13.

Egli, T., Käppeli, O. and Fiechter, A., 1982-b. Regulatory flexibility of methyltrophic yeasts in chemostat cultures: simultaneous assimilation of glucose and methanol at a fixed dilution rate. Archives of Microbiology, 131: 1-7.

Egli, T., Lindley, N. D. and Quayle, J. R., 1983. Regulation of enzyme synthesis and variation of residual methanol concentration during carbon-limited growth of *Kloeckera* sp. 2201 on mixtures of methanol and glucose. Journal of General Microbiology, 129: 1269-1281.

Egli, T., Bosshard, C. and Hamer, G., 1986. Simultaneous utilization of methanol-glucose mixtures by *Hansenula polymorpha* in chemostat: influence of dilution rate and mixture composition on utilization pattern. Biotechnology and Bioengineering, 28, 1735-1741.

Egrie, J., 1990. The cloning and production of recombinant human erythropoietin. Pharmacotherapy, 10, 3S-8S.

Elliott, S., Giffin, J., Suggs, S., Lau, E.P.L., Banks A.R., 1989. Secretion of glycosylated human erythropoietin from yeast directed by the α-factor leader region. Gene, 79, 167-180.

Elliott, S., Chang, D., Delorme, E., Eris, T., Lorenzini, T., 2004.Structural requirements for additional *N*-linked carbohydrate on recombinant human erythropoietin. Journal of Biological Chemistry, 279 (16), 16854-16862.

Ellis, S.B., Brust, P.F., Koutz, P.J., Waters, A.F., Harpold, M.M. and Gingeras, T. R., 1985. Isolation of alcohol oxidase and two other methanol regulatable genes from the yeast *Pichia pastoris*. Molecular and Cellular Biology, 5 (5), 1111-1121.

Emri, T., Molnar, Z., Pusztahelyi, T., Rosen, S., Pocsi, I., 2004. Effect of Vitamin E on Autolysis and Sporulation of Aspergillus nidulans. Applied Biochemistry and Biotechnology, 118, 337-348.

Fernandez, J.M. and Hoeffler, J.P., 1999. Gene expression systems: using nature for the art of expression. Academic Pres Inc, San Diego.

Fiaux, J., Çakar, Z.P., Sonderegger, M., Wuthrich, K., Szyperski, T., Sauer, U., 2003. Metabolic-flux profiling of the yeasts *Saccharomyces cerevisiae* and *Pichia stipitis*. Eukaryotic Cell, 2, 170–180.

Förster, J., Gombert, A.K., Nielsen, J., 2002. A functional genomics approach using metabolomics and in silico pathway analysis. Biotechnology and Bioengineering, 79, 703–712.

Förster, J., Famili, I., Fu, P., Palson, B.O., Nielsen, J., 2003. Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. Genome Research, 13, 244–253.

Franzen, C.J., 2003. Metabolic flux analysis of RQ-controlled microaerobic ethanol production by *Saccharomyces cerevisiae*. Yeast, 20 (2), 117-132.

Fredlund, E., Blank, L.M., Sauer, U., Schnurer, J., Passoth, V., 2004-a. Oxygen and glucose dependent regulation of central carbon metabolism in *Pichia anomala*. Applied and Environmental Microbiology, 70, 5905–5911.

Fredlund, E., Broberg, A., Boysen, M.E., Kenne, L., Schnurer, J., 2004-b. Metabolite profiles of the biocontrol yeast *Pichia anomala* J121 grown under oxygen limitation. Applied Microbiology and Biotechnology, 64, 403–409.

Freyre, F.M., Vazquez, J.E., Ayala, M., Canaan-Haden, L., Bell, H., Rodriguez, I., Gonzalez, A., Cintado, A., Gavilondo, J.V., 2000. Very high expression of an anti-carcinoembryonic antigen single chain Fv antibody fragment in the yeast *Pichia pastoris.* Journal of Biotechnology, 76, 157–163.

Gao, Z.H., Ackman, R.G., 1995. Determination of Vitamin-K_1 in Canolaoils by High-Performance Liquid-Chromatography with Menaquinone-4 as an Internal Standard. Food Research International, 28, 61-69.

Gemma, V., Mercedes, M., Jose, A., 2007. Optimization of integrated biodiesel production. Part 1A Study of the Biodiesel Purity and Yield. Bioresource Technology, 98, 1724-1733.

Gemmill, T.R., Trimble, R.B., 1999. Overview of *N*- and *O*-linked oligosaccharide structures found in various yeast species. Biochimica Et Biophysica Acta-General Subjects, 1426, 227-237.

Georgiou, G. and Valax, P., 1996. Expression of correctly folded proteins in *Escherichia coli*. Current Opinion in Biotechnology, 7(2): 190-197.

Gilbert J., http://www.biologymad.com, Last accessed: June 2008

Glazer, A. N., Nikaido, H., 1995. Microbial Biotechnology: Fundamentals of Applied Biotechnology, W.H. Freeman and Company, USA.

Gombert, A.K., Nielsen, J., 2000. Mathematical modeling of metabolism. Current Opinion in Biotechnology, 11, 180-186.

Gonzalez, R., Andrews, B.A., Molitor, J., Asenjo, J.A., 2003. Metabolic analysis of the synthesis of high levels of intracellular human SOD in *Saccharomyces cerevisiae* rhSOD 2060 411 SGA122. Biotechnology and Bioengineering, 82 (2), 152-169.

Gorisch, H., 1988. Drop dialysis - time course of salt and protein exchange. Analytical Biochemistry, 173, 393-398.

Guarna, M.M., Lesnicki, G.J., Tam, B.M., Robinson, J., Radziminski, C.J., Hasenwinkle, D., Boraston, A., Jervis, E., Macgillivray, R.T.A., Turner, R.F.B., Kilburn, D.G., 1997. On-line monitoring and control of methanol concentration in

shake-flask cultures of *Pichia pastoris*. Biotechnology and Bioengineering, 56:297–286.

Guile, G.R., Rudd, P.M., Wing, D.R., Prime, S.B., Dwek,R.A., 1996. A rapid high-resolution high performance liquid chromatographic method for separating glycan mixtures and analysing oligosaccharide profiles. Analytical Biochemistry, 240, 210-226.

Hamilton, S.R., Davidson, R.C., Sethuraman, N., Nett, J.H., Jiang, Y.W., Rios, S., Bobrowicz, P., Stadheim, T.A., Li, H.J., Choi, B.K., Hopkins, D., Wischnewski, H., Roser J., Mitchell T., Strawbridge R.R., Hoopes J., Wildt S., Gerngross T.U., 2006. Humanization of yeast to produce complex terminally sialylated glycoproteins. Science, 313, 1441-1443.

Harvey, D.J., 1993. Quantitative aspects of the matrix-assisted laser-desorption mass-spectrometry of complex oligosaccharides. Rapid Communications in Mass Spectrometry, 7, 614-619.

Harvey, D.J., 1999. Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates. Mass Spectrometry Reviews, 18, 349-450.

Harvey, D.J., 2005. Structural determination of N-linked glycans by matrix-assisted laser desorption/ionization and electrospray ionization mass spectrometry. Proteomics, 5, 1774-1786.

Heinonen, M., Valsta, L., Anttolainen, M., Ovaskainen, M. L., Nen, L.H., Mutanen, M., 1997. Comparisons between analyzed and calculated food composition data: Carotenoids, Retinoids, Tocopherols, Tocotrienols, Fat,Fatty Acids, and Sterols. Journal of Food Composition and Analysis, 10, 3-13.

Hellwig, S., Emde, F., Raven, N.P.G., Henke, M., Van der Logt, P., Fischer, R., 2001. Analysis of single chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentation. Biotechnology and Bioengineering, 74:344–352.

Herwig, C., von Stockar, U., 2002. A small metabolic flux model to identify transient metabolic regulations in *Saccharomyces cerevisiae*. Bioprocess and Biosystems Engineering, 24 (6), 395-403.

Higuchi, M., Oheda, M., Kuboniwa, H., Tomonoh, K., Shimonaka, Y., Ochi, N., 1992. Role of sugar chains in the expression of the biological-activity of human erythropoietin. Journal of Biological Chemistry, 267 (11): 7703-7709.

Hirose, M., Kameyama, S., Ohi, H. , 2002. Characterization of N-linked oligosaccharides attached to recombinant human antithrombin expressed in the yeast *Pichia pastoris*. Yeast, 19, 1191–1202.

Hohenblum, H., Gasser, B., Maurer, M., Borth, N., Mattanovich, D., 2004. Effects of gene dosage, promoters, and substrates on unfolded protein stress of recombinant *Pichia pastoris*. Biotechnology and Bioengineering, 85(4), 367–375.

Hong, F., Meinander, N.Q., Jonsson, L.J., 2002. Fermentation strategies for improved heterologous expression of laccase in *Pichia pastoris*. Biotechnology and Bioengineering, 79(4): 438-449.

Hutchens, T.W. and Yip, T.T., 1990. Differential interaction of peptides and protein surface structures with free metal ions and surface-immobilized metal ions. Journal of Chromatography, 500, 531–42.

Ileri, N., Çalık P., 2006. Effects of pH strategy on endo- and exo- metabolome profiles and sodium potassium hydrogen ports of beta-lactamase producing *Bacillus licheniformis.* Biotechnology Progress. 22(2), 411-419.

Inan, M., Chiruvolu, V., Eskridge, K.M., Vlasuk, G.P., Dickerson, K., Brown, S., Meagher, M.M., 1999. Optimization of temperature-glycerol-pH conditions for a fedbatch fermentation process for recombinant hookworm (*Ancylostoma caninum*) anticoagulant peptide (AcAP-5) production by *Pichia pastoris*. Enzyme and Microbial Technology, 24(7): 438-445.

Inan, M., Meagher, M.M., 2001-a. The effect of ethanol and acetone on protein expression in *Pichia pastoris*. Journal of Bioscience and Bioengineering, 92, 337-341.

Inan, M., Meagher, M.M., 2001-b. Non-repressing carbon sources for alcohol oxidase (AOX1) promoter of *Pichia pastoris*. Journal of Bioscience and Bioengineering, 92, 585-589.

Invitrogen (2002). *Pichia* fermentation process guidelines. www.invitrogen.com. Last accessed: November 2008.

Jacobs, K., Shoemaker, C., Rudersdorf, R., Neill, S.D., Kaufman, R.J., Mufson, A., Seehra, J., Jones, S.S., Hewick, R., Fritsch, E.F., 1985. Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature, 313, (6005):806-10.

Jacobson, L.O., Goldwasser, E., Fried, W., 1957. Role of the kidney in erythropoiesis. Nature (London), 179, 633-634.

Jahic, M., Rotticci-Mulder, J.C., Martinelle, M., Hult, K., Enfors, S-O., 2002. Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. Bioprocess and Biosystems Engineering, 24, 385-393.

Jahic, M., Gustavsson, M., Jansen, A. K., Martinelle, M., Enfors, S.O., 2003a. Analysis and control of proteolysis of a fusion protein in *Pichia pastoris* fed-batch processes. Journal of Biotechnology, 102: 45-53.

Jahic, M., Wallberg, F., Bollok, M., Garcia, P., Enfors, S.O., 2003b. Temperature limited fed-batch technique for control of proteolysis in *Pichia pastoris* bioreactor cultures. Microbial Cell Factories, 2: 6.

Jahic, M., Viede, A., Charoenrat, T., Teeri, T., Enfors, S.O., 2006. Process technology for production and recovery of heterologous proteins with *Pichia pastoris.* Biotechnology Progress, 22, 1465-1473.

Jelkmann, W., 1992. Erythropoietin: structure, control of production, and function. Physiological Reviews, 72, 449– 489.

Jenny, R.J., Mann, K.G., Lundblad, R.L., 2003. A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. Protein Expression and Purification, 31, 1–11.

Jungo, C., Rerat, C., Marison, I.W., von Stockar, U., 2006. Quantitative characterization of the regulation of the synthesis of alcohol oxidase and of the expression of ecombinant avidin in a *Pichia pastoris* Mut^+ strain. Enzyme and Microbial Technology, 39, 936-944.

Jungo, C., Schenk, J., Pasquier, M., Marison, I.W., von Stockar, U., 2007. A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. Journal of Biotechnology, 131, 57-66.

Karas, M., Hillenkamp F., 1988. Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Analytical Chemistry, 60, 2299-2301.

Katakuga, Y., Zhang, W., Zhuang, G., Omasa, T., Kishimoto, M., Goto, Y., Suga, K.I., 1998. Effect of methanol concentration on the production of human beta-2-glycoprotein I domain V by a recombinant *Pichia pastoris:* A simple system for control of methanol concentration using a semiconductor gas sensor. Journal of Fermentation and Bioengineering, 86:482–487.

Kazusa DNA Research Institute, http://www.kazusa.or.jp/codon/, Last accessed: June 2008.

Kim, Y.K., Shin, H.S., Tomiya, N., Lee, Y.C., Betenbaugh, M.J., Cha, H.J., 2005. Production and N-glycan analysis of secreted human erythropoietin glycoprotein in stably transfected Drosophila S2 cells. Biotechnology and Bioengineering, 92, 452-461.

Kirk, R.E. and Othmer, D.F., 1994. Encyclopedia of Chemical Technology, 4th ed., The Interscience Encyclopedia Inc., New York.

Klug, W. S., Cummings, M. R., Spencer, C. A., 2006. Concepts of genetics, Eight Edition, Pearson Prentice Hall, New Jersey.

Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T., Ohyama, M., Tomomitsu, K., 2000-a. High level secretion of recombinant human serum albumin by fedbatch fermentation of the methylotrophic yeast, *Pichia pastoris*, based on optimal methanol feeding strategy. Journal of Bioscience and Bioengineering, 90:280–288.

Kobayashi, K., Kuwae, S., Ohya, T., Ohda, T., Ohyama, M., Tomomitsu, K., 2000-b. High level secretion of recombinant human serum albumin by fedbatch fermentation of the methylotrophic yeast, *Pichia pastoris*, with minimal protease

production and activation. Journal of Bioscience and Bioengineering, 89(1): 55-61.

Koganesawa, N., Aizava, T., Shimojo, H., et al. 2002. Expression and purification of a small cytokine growth-blocking peptide from armyworm Pseudaletia separata by an optimized fermentation method using the methylotrophic yeast *Pichia pastoris*. Protein Expression and Purification, 25(3): 416–425.

Koutz, P., Davis, G. R., Stillman, C., Barringer, K., Cregg, J., Thill, G., 1989. Structural comparison of the *Pichia pastoris* alcohol oxidase genes. Yeast, 5 (3), 167-177.

Kresnowati, M.T.A.P., van Winden, W.A., van Gulik, W.M., Heijnen, J.J. 2008. Dynamic In Vivo metabolome response of *Saccharomyces cerevisiae* to a stepwise perturbation of the ATP requirement for Benzoate Export. Biotechnology and Bioengineering, 99, 421- 441.

Kupesulik, B. and Sevella, B., 2005. Optimization of specific product formation rate by statistical and formal kinetic model description of an HSA producing *Pichia pastoris* MutS strain. Chemical and Biochemical Engineering Quarterly, 19 (1): 99-108.

Küster, B., Wheeler, S.F., Hunter, A.P., Dwek, R.A., Harvey, D.J., 1997. Sequencing of N-linked oligosaccharides directly from protein gels: In-gel deglycosylation followed by matrix-assisted laser desorption/ionization mass spectrometry and normal-phase high-performance liquid chromatography. Analytical Biochemistry, 250, 82-101.

Laemmli, U. K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, 227, 680-685.

Lai, P.H., Everett, R., Wang, F.F., Arakawa, T., Goldwasser, E., 1986. Structural characterization of human erythropoietin. Journal of Biological Chemistry, 216: 3116-3121.

Large, P.J., 1986. Degradation of organic nitrogen compounds by yeasts. Yeast, 2, 1-34.

Lasne, F., de Ceaurriz, J., 2000. Recombinant erythropoietin in urine. Nature, 405, 635.

Law, M.L., Cai, G.Y., Lin, F.K., Wei, Q., Huang, S.Z., Hartz J.H., Morse, H., Lin, C.H., Jones, C., Kao, F.T., 1986. Chromosomal assignment of the human erythropoietin gene and its DNA polymorphism. Proceedings of the National Academy of Sciences USA, 83(18):6920-4.

Lee, B., Yurimoto, H., Sakai, Y. and Kato, N., 2002. Physiological role of the glutathione-dependent formaldehyde dehydrogenase in the methylotrophic yeast *Candida boidinii*. Microbiology, 148, 2697-2704.

Leehuang, S., 1984. Cloning and Expression of Human Erythropoietin cDNA in *Escherichia coli*. Proceedings of The National Academy of Sciences USA-Biological Sciences, 81 (9), 2708-2712.

Lehninger, A.L., Nelson, D.L., Cox, M.M., 1993. Principles of Biochemistry, 2nd edition, Worth Publishers Inc., New York.

Li, A., Crimmins, D.L., Luo, Q., Hartupee, J., Landt, Y., Ladenson, J.H., Wilson, D., Anant, S., Dieckgraefe, B.K. (2003). Expression of a novel regenerating gene product, Reg IV, by high density fermentation in *Pichia pastoris* : production, purification, and characterization. Protein Expression and Purification, 31(2): 197-206.

Li, Z., Xiong, F., Lin, Q., D'Anjou, M., Daugulis, A.J., Yang, D.S.C., Hew, C.L., 2001. Low-temperature increases the yield of biologically active herring antifreeze protein in *Pichia pastoris*. Protein Expression and Purification, 21: 438-445.

Lin, F-K., Suggs, S., Lin, C-H., Browne, J.K., Smalling, R., Egrie, J.C., Chen, K.K., Fox, G.M., Martin, F., Stabinsky, Z., Badrawi, S.M., Lai, P-H., Goldwasser, E., 1985. Cloning and expression of the human erythropoietin gene. Proceedings of the National Academy of Sciences USA, 82(22):7580-4.

Loewen, M.C., Liu, X., Davies, P.L. and Daugulis, A.J., 1997. Biosynthetic production of type II fish antifreeze protein: fermentation by *Pichia pastoris*. Applied Microbiol Biotechnology, 48(4), 480-486.

Lopez-Soto-Yarritu, P., Diez-Masa, J.C., de Frutos, M., Cifuentes, A., 2002. Comparison of different capillary electrophoresis methods for analysis of recombinant erythropoietin glycoforms. Journal of Separation Science, 25: 1112-1118.

Lüers, G.H., Advani, R., Wenzel, T. and Subramani, S., 1998. The *Pichia pastoris* dihydroxyaceton kinase is a PTS1-containing, but cytosolic, protein that is essential for growth on methanol. Yeast, 14, 759-771.

Maaheimo, H., Fiaux, J., Cakar, Z.P., Bailey, J.E., Sauer, U., Szyperski, T., 2001. Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional ^{13}C labeling of common amino acids. European Journal of Biochemistry, 268, 2464-2479.

Macauley-Patrick, S., Fazenda, M.L., McNeil, B., Harvey, L.M., 2005. Heterologous protein production using the *Pichia pastoris* expression system. Yeast, 22, 249–270.

Maiese K., Li, F., Chong, Z.Z., 2005. New Avenues of Exploration for Erythropoietin, Journal of the American Medical Association, 293: 90-95.

Mathews, C.K., van Holde, K.E., 1996. Biochemistry, 2nd ed., Benjamin-Cummings Pub Co., USA

McGrew, J.T., Leiske, D., Dell, B., Klinke, R., Krasts, D., Wee, S., Abbott, N., Armitage, R. and Harrington, K., 1997. Expression of trimeric CD41 ligand in *Pichia pastoris*: use of a rapid method to detect high-level expressing transformants. Gene, 187(2): 193-200.

Menendez, J., Valdes, I., Cabrera, N., 2003. The ICL1 gene of *Pichia pastoris*, transcriptional regulation and use of its promoter. Yeast, 20, 1097-1108.

Minning, S., Serrano, A., Ferrer, P., Sola, C., Schmid, R.D., Valero, F., 2001. Optimization of high-level production of *Rhizopus oryzae* lipase in *Pichia pastoris:* Enzyme production using plasmid pPICZ-alpha-AROL. Journal of Biotechnology, 86:59–70.

Miyake, T., Kung, C., Goldwasser, E., 1977. Purification of human erythropoietin. Journal of Biological Chemistry, 252(15):5558-64.

Moon, S.H., Parulekar, S.J., 1991. A Parametric study of protease production in batch and fed-batch cultures of *Bacillus firmus*. Biotechnology and Bioengineering, 37, 467-483.

Morelle, W., Michalski, J.C., 2005. The mass spsectrometric analysis of glycoproteins and their glycan structures. Current Analytical Chemistry, 1, 29-57.

Murasugi, A., Tohma-Aiba, Y., Asami, Y., 2000. Production of recombinant human midkine in yeast, *Pichia pastoris:* Induction by methanol in high cell density fermentation. Journal of Bioscience and Bioengineering, 90:395–399.

Nagao, M., Inoue, K., Moon, S.K., Masuda, S., Takagi, H., Udaka, S., Sasaki, R., 1997. Secretory production of erythropoietin and the extracellular domain of the erythropoietin receptor by Bacillus brevis: affinity purification and characterization. Bioscience, Biotechnology, and Biochemistry, 61 (4), 670-4.

National Library of Medicine, http://www.ncbi.nlm.nih.gov/BLAST/, Last accessed: September 2008.

Naven, T.J.P., Harvey, D.J., 1996. Effect of structure on the signal strength of oligosaccharides in matrix-assisted laser desorption/ionization mass spectrometry on time-of-flight and magnetic sector instruments. Rapid Communications in Mass Spectrometry, 10, 1361-1366.

Nevoigt, E. and Stahl, U., 1997. Osmoregulation and glycerol metabolism in the yeast *Saccharomyces cerevisiae*. FEMS Microbiology Reviews, 21, 231-241.

New England Biolabs Inc., http://tools.neb.com/NEBcutter2/index.php, Last accessed: June 2008.

Nielsen, J., Villadsen, J., Liden, G., 2003. Bioreaction Engineering Principles. Kluwer Academic Publishers, 2nd ed.

Nielsen, J., and Jewett, M.C., 2008. Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. FEMS Yeast Research, 8, 122-131.

Nilsson, J., Stahl, S., Lundeberg, J., Uhlén. M., Nygren. P.A., 1997. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. Protein Expression and Purification, 11, 1–16.

Nissen, T.L., Schulze, U., Nielsen, J., Villadsen, J., 1997. Flux distribution in anaerobic, glucose-limited continuous cultures of *Saccharomyces cerevisiae*. Microbiology, 143, 203-218.

Pais, J.M., Varas, L., Valdes, J., Cabello, C., Rodriguez, L., Mansur, M., 2003. Modeling of mini-proinsulin production in *Pichia pastoris* using the AOX promoter. Biotechnology Letters, 25, 251-255.

Papac, D.I., Wong, A., Jones, A.J.S., 1996. Analysis of acidic oligosaccharides and glycopeptides by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Analytical Chemistry, 68, 3215-3223.

Pitkanen, J.P., Aristidou, A., Salusjarvi, L., Ruohonen, L., Penttila, M., 2003. Metabolic flux analysis of xylose metabolism in recombinant *Saccharomyces cerevisiae* using continuous culture. Metabolic Engineering, 5 (1), 16-31.

Plantz, B.A., Nickerson, K., Kachman, S.D., Schlegel, V.L., 2007. Evaluation of Metals in a Defined Medium for *Pichia pastoris* Expressing Recombinant *β*-Galactosidase. Biotechnology Progress, 23, 687-692.

Protein Databank, http://www.rcsb.org/pdb/explore/explore.do?structureId =1BUY, Last accessed: September 2008.

Raghevendran, V., Gombert, A.K., Christensen, B., Kotter, P., Nielsen, J., 2004. Phenotypic characterization of glucose repression mutants of *Saccharomyces cerevisiae* using experiments with C-13-labelled glucose. Yeast, 21 (9), 769-779.

Ramon, R., Ferrer, P., Valero, F., 2007. Sorbitol co-feeding reduces metabolic burden caused by theoverexpression of a *Rhizopus oryzae* lipase in *Pichia pastoris.* Journal of Biotechnology, 130: 39-46.

Recny, M.A, Scoble, H.A, Kim, Y., 1987. Structure characterization of natural urinary and recombinant DNA-derived erythropoietin. Journal of Biological Chemistry, 262: 17156-17163.

Roe, S., 2001. Protein purification techniques, 2nd edition. Oxford University Pres, Oxford.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor, USA.

Sarramegna, V., Demange, P., Milon, A., Talmont, F., 2002. Optimizing functional versus total expression of the human m-opioid receptor in *Pichia pastoris*. Protein Expression and Purification, 24: 212-220.

Sasaki, H., Bothner, B., Dell, A., Fukuda, M., 1987. Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. Journal of Biological Chemistry, 262:12059-12076.

Sasaki, H., Ochi, N., Dell, A., Fukuda, M., 1988. Site-specific glycosylation of human recombinant erythropoietin: Analysis of glycopeptides or peptides at each glycosylation site by fast atom bombardment mass spectrometry. Biochemistry, 27: 8618-9626.

Schenk, J., Balazs, K., Jungo, C., Urfer, J., Wegman, C., Zocchi, A., Marison, I.W., von Stockar, U., 2008. Influence of specific growth rate on specific productivity and glycosylation of a recombinant avidin produced by a *Pichia pastoris* Mut^+ strain. Biotechnology and Bioengineering, 99, 368-376.

Schmidt, F.R. 2004. Recombinant expression systems in the pharmaceutical industry. Applied Microbiology and Biotechnology, 65, 363-372.

Shen, S., Sulter, G., Jeffries, T.W., Cregg, J.M., 1998. A strong nitrogen sourceregulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. Gene, 216, 93-102.

Shi, X., Karkut, T., Chamankhah, M., Alting-Mees, M., Hemmingsen, S. M., Hege dus, D., 2003. Optimal conditions for the expression of a single-chain antibody (scFv) gene in *Pichia pastoris*. Protein Expression and Purification, 28: 321-330.

Shuler, M.L. and Kargı, F., 2002. Bioprocess Engineering: Basic Concepts, 2nd Ed., Prentice Hall Inc., USA.

Sinha, J., Plantz, B.A., Zhang, W., Gouthro, M., Schlegel, V.L., Liu, C.P., Meagher, M.M., 2003. Improved production of recombinant ovine interferon-ô by Mut+ strain of *Pichia pastoris* using an optimized methanol feed profile. Biotechnology Progress, 19, 794-802.

Skibeli, V., Lie, G.N., Torjesen, P., 2001. Sugar profiling proves that human serum erythropoietin differs from recombinant human erythropoietin. Blood, 98 (13): 3626-3634.

Smith, A., 1995. Gene Expression in Recombinant Microorganisms, 1st Ed., Marcel Dekker Inc., USA.

Sola, A., Maaheimo, H., Ylonen, K., Ferrer, P., Szyperski, T., 2004. Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*. European Journal of Biochemistry. 271 (12), 2462-2470.

Sola, A., Jouhten, P., Maaheimo, H., Sanchez-Ferrando, F., Szyperski, T., Ferrer, P., 2007. Metabolic flux profiling of *Pichia pastoris* grown on glycerol/methanol mixtures in chemostat cultures at low and high dilution rates, Microbiology, 153, 281–290.

Sreekrishna, K. and Kropp, K.E., 1996. *Pichia pastoris*. Nonconventional yeasts in biotechnology. Wolf, K. Berlin, Springer**,** 203-253.

Sreekrishna, K., Brankamp, R.G., Kroop, K.E., Blankenship, D.T., Tsay, J.T., Smith, P.L., Wierschke, J.D., Subramaniam, A., Birkenberger, L.A., 1997. Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. Gene, 190, 55-62.

Stanley, S.M.R., Poljak, A., 2003. Matrix-assited laser-desorption time-of-flight ionization and high performance liquid chromatography-electrospray ionization mass spectral analyses of two glycosylated recombinant epoetins. Journal of Chromatography B, 785, 205-218.

Stratton, J., Chiruvolu, V., Meagher, M., 1998. High cell-density fermentation. Methods in Molecular Biology. D.R. Higgins, J. M. C., Humana Press. *Pichia* Protocols, 107-120.

Sttübiger, G., Marchetti, M., Nagano, M., Reichel, C., Gmeiner, G., Allmaier, G., 2005. Characterization of intact recombinant human erythropoietin applied in

doping by means of planar gel electrophoretic techniques and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Communications in Mass Spectrometry, 19, 728-742.

Takeuchi, M., Takasaki, S., Miyazaki, H., Kato, T., Hoshi, S., Kochibe, N., Kobata, A., 1988. Comparative study of the asparagine-linked sugar chains of human erythropoietin purified from urine and the culture medium of recombinant Chinese hamster ovary cells. Journal of Biological Chemistry, 263: 3657-3663.

Takeuchi, M., Inoue, N., Strickland, T.W., Kubota, M., Wada, M., Shimizu, R., Hoshi, S., Kozutsumi, H., Takasaki, S., Kobata, A., 1989. Relationship between sugar chain structure and biological-activity of recombinant human erythropoietin produced in Chinese hamster ovary cells. Proceedings of the National Academy of Sciences USA, 86 (20): 7819-7822.

Takeuchi, M., Takasaki, S., Shimada, M., Kobata, A., 1990. Role of sugar chains in the *in vitro* biological-activity of human erythropoietin produced in Chinese hamster ovary cells. Journal of Biological Chemistry, 265 (21): 12127-12130.

Tarentino A.L., Gomez C.M., Plummer T.H., 1985. Deglycosylation of asparagine-linked glycans by peptide: N-glycosidase F. Biochemistry, 24, 4665-4671.

Terpe, K., 2003. Overview of tag protein fusions: From molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology, 60, 523–533.

Thompson, J.C., He, B.B., 2006. Characterization of crude glycerol from biodiesel production from multiple feedstocks. Applied Engineering in Agriculture, 22, 261-265.

Thorpe, E.D., d'Anjou, M.C., Daugulis, A.J., 1999. Sorbitol as a non-repressing carbon source for fed-batch fermentation of recombinant *Pichia pastoris*. Biotechnology Letters, 21, 669–672.

Trentmann, O., Khatri, N.K., Hoffmann, F., 2004. Reduced oxygen supply increases process stability and product yield with recombinant *Pichia pastoris.* Biotechnology Progress, 20, 1766-1775.

Trinh, L.B., Phue, J.N., Shiloach, J., 2003. Effect of methanol feding strategies on production and yield of recombinant mouse endostatin from *Pichia pastoris.* Biotechnology and Bioengineering, 82, 438-444.

Tschopp, J.F., Brust, P.F., Cregg, J.M., Stillman, C.A., Gingeras, T.R., 1987. Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*. Nucleic Acids Research, 15(9): 3859-3876.

Tsuda, E., Goto, M., Murakami, A., Akai, K., Ueda, M., Kawanishi, G., Takahashi, N., Sasaki, R., Endo, S., Arata, Y., 1988. Comparative structural study of N-linked oligosaccharides of urinary and recombinant erythropoietins. Biochemistry, 27: 5646-5654.

Tsuda, E., Kawanihi, G., Ueda,M., Masuda, S., Sasaki, R., 1990. The role of carbohydrate in recombinant human erythropoietin. European Journal of Biochemistry, 188 (2): 405-411.

van Gulik, W.M., Heijnen, J.J., 1995. A metabolic network stoichiometry analysis of microbial-growth and product formation. Biotechnology and Bioengineering, 48 (6), 681-698.

Veenhuis, M., Vandijken, J.P., Harder, W., 1983. The significance of peroxisomes in the metabolism of one-carbon compounds in yeasts. Advances in Microbial Physiology, 24, 1-82.

Villadsen, J., Nielsen, J., 1990. Modeling of fermentation kinetics, in Proceedings of European Congress Biotechnology, 259-266, Copenhagen.

Wahlbom, C.F., Eliasson, A., Hahn-Hagerdal, B., 2001. Intracellular fluxes in a recombinant xylose-utilizing *Saccharomyces cerevisiae* cultivated anaerobically at different dilution rates and feed concentrations. Biotechnology and Bioengineering, 72 (3), 289-296.

Walker, G.M., 1998. Yeast: Physiology and Biotechnology. John Wiley & Sons Inc., New York.

Waterham, H.R., Digan, M.E., Koutz, P.J., Lair, S.V., Cregg, J.M., 1997. Isolation of the *Pichia pastoris* glyceraldehyde-3-phosphate dehydrogenase gene and regulation and use of its promoter. Gene, 186, 37-44.

Wegner, E.H., 1983. Biochemical conversions by yeast fermentation at high-cell densities. U.S. Patent 4, 414,329. US, Phillips Petroleum Company.

Wegner G., 1990. Emerging applications of the methylotrophic yeasts. FEMS Microbiology Reviews, 7: 279-283.

Weigand, W.A., Lim, H.C., Creagan, C.C., Mohlar, R.D., 1979. Optimization of repeated fed-batch reactor for maximum cell productivity. Biotechnology Bioengineering Symposium 9, 335-348.

White, C.E., Kempi, N.M., Komives, E.A., 1994. Expression of highly disulfide-bonded proteins in *Pichia pastoris*. Structure, 2 (11), 1003–1005.

Whittaker, M.M. and Whittaker, J.W., 2000. Expression of recombinant galactose oxidase by *Pichia pastoris*. Protein Expression and Purification, 20: 105-111.

Wiechert, W., 2002. Modeling and simulation: tools for metabolic engineering. Journal of Biotechnology, 94 (1), 37-63.

Xie, J., Zhou, Q., Du, P., Gan, R. and Ye, Q., 2005. Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostatin production. Enzyme and Microbial Technology, 36, 210-216.

Yang, J., Zhou, X.S., Zhang, Y.X., 2004. Improvement of recombinant hirudin production by controlling NH4+ concentration in *Pichia pastoris* fermentation. Biotechnology Letters, 26:1013-1017.

Yip, T.T., Nakagawa, Y., Porath, J., 1989. Evaluation of the interaction of peptides with Cu(II), Ni(II), and Zn(II) by high-performance immobilized metal ion affinity chromatography. Analytical Biochemistry, 183, 159–71.

Zanjani, E.D., Poster, J., Borlington, H., Mann, L.I., Wasserman L.R., 1977. Liver as the primary site of erythropoietin formation in the fetus. Journal of Laboratory and Clinical Medicine, 89, 640.

Zhang, W., Bevins, M.A., Plantz, B.A., Smith, L.A., 2000a. Modeling *Pichia pastoris g*rowth on methanol and optimizing the production of a recombinant protein, the heavy-chain fragment C of Botolinum Neurotoxin, Serotype A. Biotechnology and Bioengineering, 70 (1), 1-8.

Zhang, W., Inan, M. and Meagher, M.M. 2000-b. Fermentation strategies for recombinant protein expression in the methylotrophic yeast *Pichia pastoris*. Biotechnology and Bioprocess Engineering, 5, 275-287.

Zhang, W.H., Potter, K.J.H., Plantz, B.A., Schlegel, V.L., Smith, L.A., Meagher, M.M., 2003-a. *Pichia pastoris* fermentation with mixed-feeds of glycerol and methanol: growth kinetics and production improvement. Journal of Industrial Microbiology and Biotechnology, 30, 210-215.

Zhang, H.M., Shimizu, K., Yao, S.J., 2003-b. Metabolic flux analysis of *Saccharomyces cerevisiae* grown on glucose, glycerol or acetate by C-13-labeling experiments. Biochemical Engineering Journal, 16 (3), 211-220.

Zhang, W., Sinha, J., Smith, L.A., Inan, M., Meagher, M.M., 2005. Maximization of production of secreted recombinant proteins in *Pichia pastoris* fed-batch fermentation. Biotechnology Progress, 21 (2), 386–393.

APPENDIX A

CHEMICALS AND SUPPLIERS

Chemical	**Supplier**
Acetic acid	BDH
Acetonitrile, HPLC grade	Sigma
Agarose (electrophoresis grade)	Melford
4-Aminoantipyrine (4-AAP)	Sigma
Ammonia solution, 25% (NH_3OH)	Sigma
Ammonium persulphate (APS)	Sigma
Ammonium sulphate	BDH
Antibody, Monoclonal Anti-human EPO	R&D Systems
Antifoam Y-30 emulsion	Sigma
Antimouse IgG, Horseradish peroxidase linked whole antibody	Amersham
AOX from *P. pastoris*	Sigma
Bacto peptone	Difco
Bacto yeast extract	Difco
D-Biotin	Sigma
Borax ($Na_2B_4O_7.10H_2O$) - Sodium Borate	Merck
Boric Acid (H_3BO_3)	Merck
Briallant Blue R250	Sigma
Bromophenol blue	Sigma
Calcium chloride 2-hydrate	BDH, Merck
Calcium sulfate	Merck
Citric acid monohydrate	Merck
Cobalt Chloride hexahydrate ($CoCl_2.6H_2O$)	Merck
Copper Sulphate Pentahydrate ($CuSO_4.5H_2O$)	Merck
Cytochrome *c*	Sigma
DNA ladder – Hyperladder I	Bioline
dNTPs	Bioline
Ethanol	BDH
Ethanol, HPLC grade	Sigma

Ethidium bromide	Sigma
Ethylenediaminetetraacetic acid disodium salt dihydrate (EDTA)	Sigma
Ferric chloride hexahydrate ($FeCl_3.6H_2O$)	Merck
ferrous sulphate heptahydrate ($FeSO_4.7H_2O$)	Merck
Formaldehyde (37%)	J.T. Baker
Glass beads	Sigma
D-Glucose	BDH, Merck
Glycerol	Sigma
Glycine	Sigma
GlycoBlue	Ambion
Horseradish peroxidase (HRP)	Sigma
Hydrochloric acid	BDH, Merck
isopropanol	BDH
Kanamycin	Sigma
Lithium chloride	BDH, Merck
Magnesium chloride 6-hydrate	BDH
Magnesium Sulphate heptahydrate ($MgSO_4.7H_2O$)	Merck
Manganese Sulphate, Monohydrate ($MnSO_4.H_2O$)	Merck
$(Man)_9$-$(GlcNAc)_2$	Sigma
2-Mercaptoethanol	Sigma
MES	Sigma
Methanol	BDH
N,N,N',N'- Tetramethylethylene diamine (TEMED)	Sigma
Non-fat milk powder	Marvel, Fluka
Phenol-4-sulfonic acid (PSA)	Sigma
PNGaseF (N-Glycanase)	Glyko
Polyethylene glycol, 3350	Sigma
Potassium acetate	BDH
Potassium carbonate	Sigma
Potassium chloride	BDH
Potassium dihydrogen orthophoshate	BDH, Merck
di-Potassium hydrogen orthophosphate	BDH, Merck
Potassium hydroxide	Merck
Potassium Iodide	Merck
Prestained protein ladder	BioRad, Fermentas
Primers	MWG

Protogel ®	National Diag., UK
Restriction enzymes	Roche
RNaseA	Sigma
S-DHB, for MS analysis	Sigma
Silver nitrate	Fluka
Sinapinic acid, for MS analysis	Sigma
Sodium acetate	Sigma
Sodium chloride	BDH, Merck
Sodium dihydrogen orthophosphate, anhydrous	BDH, Merck
Sodium dodecyl suplhate (SDS)	Sigma
di-Sodium hydrogen orthophospahete, anhydrous	BDH, Merck
di-Sodium hydrophosphate heptahydrate	Merck
Sodium hydroxide	BDH, Merck
Sodium Iodide	Merck
Sodium Molybdate ($Na_2MoO_4.2H_2O$)	Merck
Sodium thiosulphate ($Na_2S_2O_3.5H_2O$)	Fluka
D-Sorbitol	Sigma
Sucrose	BDH
Sulfiric acid	Sigma
T4 DNA polymerase	Roche
Taq DNA polymerase	Roche
TFA, HPLC grade	Pierce
Trisodium citrate	BDH
Triton X-100	BDH
Trizma Base	Sigma
Tryptone	Difco
Tween 20	Sigma
Urea	BDH
Xylene cyanol FF - green to orange	BDH
Yeast nitrogen base w/o amino acids and ammonium sulphate	Sigma
Zeocin	Invitrogen
Zinc chloride	Merck
Zinc Sulphate heptahydrate ($ZnSO_4.7H_2O$)	Merck
Zwitter ion	Waters

APPENDIX B

LABORATORY EQUIPMENT AND SUPPLIERS

Equipment	**Supplier**
Autoclave	HVE-50, Hirayama (Japan)
Balances	Sartorius (Germany)
Centrifuges	5415 R, Eppendorf (Germany) Avanti J-20, Beckman Coulter (USA) Z323K, Hermle (Germany)
Deep freezers	Legaci, Revco (USA)
Freeze drier	Modulyo D, Savant (USA)
Gel documentation system	Gel Doc2000, BioRad (USA)
HPCE	Waters (USA)
HPLC	Waters (USA)
Horizontal gel-electrophoresis	BioRad (USA)
Hybridization oven	HB-1D, Techne (UK)
Incubators	GenLab (UK)
	Nüve (Turkey)
Laminar flow cabinet	Nuaire (USA)
MALDI-ToF MS	Waters, Micromass (USA)
Micropipets	Gilson
Microscope	HC, Leica
Nanodrop	Nanodrop Technologies (USA)
Orbital shakers	C25, C25KC, New Brunswick Scientific (USA)
	Certomat BS-T, BS-I, B.Braun (Germany)
PCR machine	GeneAmp-9700, Applied Biosystems (USA)
pH meter	Sartorius (Germany)
Refrigerators	Thermo Electron-Revco (USA)
Spectrophotometer	Helios alpha, ThermoSpectronics (UK)
	Ce2021, Cecil (UK)
Ultrafiltration cell	Amicon, Millipore (USA)
Vacuum evaporator	SPD111V, Savant (USA)

Vertical gel-electrophoresis	BioRad (USA)
Vortex	Labnet (USA)
Water baths	Grant (UK)
Water purification system	Millipore (USA)

APPENDIX C

BUFFERS AND STOCK SOLUTIONS

0.125 M (or 0.5 M) EDTA, pH 8.0

4.65 g (or 18.61 g) Ethylenediaminetetra acetic acid disodium salt dihydrate was was dissolved in 80 mL dH_2O. NaOH was added until EDTA was dissolved. The final pH was further adjusted to pH 8.0 and the final volume was adjusted to 100 mL. The buffer was autoclaved and stored at room temperature.

50 x TAE

242 g Tris base was dissolved in 57.1 mL acetic acid and 100 mL 0.5 M EDTA, pH8.0, made up to 1 liter with sterile dH_2O and stored at room temperature. Before use, the stock solution was diluted 1:50 with dH_2O.

1 M Tris-Cl, pH 8.0

12.1 g Tris base was dissolved in 80 mL dH_2O and the pH was adjusted to 8.0 by adding concentrated HCl. The volume was made up to 100 mL. The buffer was autoclaved and stored at room temperature.

TE Buffer, pH 8.0

1 mL of 1M Tris-Cl (pH 8.0), 200 µL of 0.5 M EDTA (pH 8.0) was added to dH_2O and the volume was made up to 100 mL. The buffer was autoclaved and stored at room temperature.

50 mM potassium acetate, pH 5.5

0.049 g potassium acetate, dissolved in 8.5 mL dH_2O, titrated with 10N glacial acetic acid to pH 5.5 and make up to 10 mL with dH_2O. Autoclaved and store at room temperature.

RNaseA stock solution

RNaseA was dissolved at a concentration of 10 mg mL^{-1} in 50 mM potassium acetate (pH 5.5) and boiled for 10 min. Stored at -20 $^{\circ}$ C.

Yeast Lysis Solution	2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl-pH8.0, 1mM Na_2EDTA. The solution was autoclaved and stored at room temperature.
6 x DNA gel-loading buffer	0.25% Bromophenol blue, 0.25% xylenecyanol FF, 40 % sucrose in dH_2O. Stored at room temperature.
3 M Sodium acetate, pH 5.2	24.6 g sodium acetate was dissolved in 80 mL dH_2O and the pH was adjusted to 5.2 with 3M acetic acid. The buffer was filter sterilized and stored at 2-8°C.
1.5 M Tris-HCl, pH 8.8	36.3 g Tris base was dissolved in 150 mL dH_2O and pH was adjusted to 8.8 with 6N HCl. The buffer was made up to 200 mL with dH_2O. The buffer was autoclaved and stored at 2-8°C.
0.5 M Tris-HCl, pH 6.8	12.1 g Tris base was dissolved in 150 mL dH_2O and pH was adjusted to 6.8 with 6N HCl. The buffer was made up to 200 mL with dH_2O. The buffer was autoclaved and stored at 2-8°C.
4 x Sample Loading Buffer for SDS-PAGE	200 mM Tris-HCl,pH 6.8; 40% glycerol; 6% SDS; 0.013% Bromophenol blue; 10% 2-mercaptoethanol. Distributed into microcentrifuge tubes and stored at -20°C.
5x SDS-PAGE Running Buffer	15 g Tris Base, 72 g glycine, 5 g SDS, dH2O to 1 liter. The buffer was stored at 2-8°C and diluted 1:5 with dH_2O prior to use.
Coomassie Staining Solution	0.25 g Coomassie Brilliant Blue R250, 45 mL methanol, 10 mL Acetic acid was added to 45 mL dH_2O.
Destaining Solution	180 mL methanol and 40 mL acetic acid was added to 180 mL dH_2O. Stored at room temperature.
10X Transfer buffer stock solution for Western blot	30.3 g Tris base and 144.1 g glycine was dissolved in dH_2O and the volume was made upto 1 liter. The buffer was stored at 2-8°C.

1X Transfer buffer for Western blot	100 mL 10X transfer buffer stock, 200 mL methanol, dH_2O to 1 liter. Prepared on the day of use and stored at 2-8°C.
10X TBS	12.11 g Tris-base, 87.66 g NaCl dissolved in 900 mL dH_2O and pH adjusted to 7.6. Stored at room temperature.
TBS-T solution	0.1% Tween-20 was added into 1xTBS solution. Prepared on the day of use.
TBS-T – Milk	5 % non-fat milk powder in TBS-T. Prepared on the day of use.
Reaction buffer for factor Xa digestion	2.42 g Tris-base, 2.92 g NaCl, 0.147 g $CaCl_2.2H_2O$ in 900 mL dH_2O. pH adjusted to 6.5 using HCl and volume was made upto 1 liter. The buffer was autoclaved and stored at 2-8°C.
Denaturation soln. for PNGaseF digestion	2% SDS, 1 M β-mercaptoethanol (comes with the enzyme)
Detergent solution for PNGaseF digestion	15% NP-40 solution (comes with the enzyme)
5x PNGaseF reaction buffer	100 m*M* sodium phosphate; 0.1% sodium azide, pH 7.5 (comes with the enzyme)
1x Equilibration / wash buffer for His-Tag purification	50 mM sodium phosphate buffer, pH 7.0; 300 mM NaCl.
1x Elution buffer for His-Tag purification	50 mM sodium phosphate buffer, pH 7.0; 300 mM NaCl; 150 mM Imidazole.
MES buffer for resin washing	20 mM MES Buffer (pH 5.0); 0.1 M NaCl.

1 M potassium phosphate, pH 6.0	56.48 g KH_2PO_4, 14.8 g K_2HPO_4 was dissolved in dH_20 and the volume made upto 500 mL. The pH was controlled. The buffer was autoclaved and stored at room temperature.
20x YNB Stock solution	17 g Yeast Nitrogen Base without amino acids, 50 g $(NH_4)_2SO_4$ was dissolved in dH_20 and the volume made upto 500 mL. The solution was autoclaved, aliquoted into 50 mL Falcon® tubes and stored at room temperature in dark.
Primary wash buffer for Southern Blot	120 g Urea, 1 g SDS, 100 mL 0.5 M Na phosphate pH 7.0, 8.7 g NaCl, 1 mL 1.0M $MgCl_2$, 2 g Blocking reagent was dissolved in sterile dH_20 and the volume made upto 1 liter. The buffer was stored at 2-8°C for upto 1 week.
20x Secondary wash stock buffer for Southern blot	121 g Tris base, 112 g NaCl was dissolved in 800 mL sterile dH_20. The pH was adjusted to 10.0 and the volume made upto 1 liter. The buffer was stored at 2-8°C.
1x Secondary wash buffer for Southern blot	20x secondary wash stock buffer was diluted 1:20 and 2 mL L^{-1} of 1 M $MgCl_2$ was added to give a final concentration of 2 mM magnesium in the buffer. This buffer should not be stored.
Depurination solution	0.25 M HCl. Stored at room temperature.
Denaturation solution for Southern blot	1.5 M NaCl, 0.5 M NaOH. Stored at room temperature.
Neutralising solution for Southern blot	1.0 M Tris base, 1.5 M NaCl. The pH was adjuted to 7.5 with HCl. Stored at room temperature.
Transfer Buffer for Southern blot (20x SSC)	175 g NaCl and 88.2 g trisodium citrate was dissolved in 800 mL dH2O. The pH was adjusted to 7.0-7.2 with concentrated citric acid. Stored at room temperature.

Buffer H (for *Eco*R I and *Xba* I)	500 mM Tris-HCl, 100 mM $MgCl_2$, 1M NaCl, 10 mM DTE, pH=7.5
Buffer A (for *Sac* I)	330 mM Tris acetate, 100 mM Mg-acetate, 660 mM K-acetate, 5 mM DTT, pH=7.9
Single-stranded carrier DNA	500 mg of high MW DNA (Sigma) was added to 100 mL TE buffer ph 8.0 and stored at -20°C. Before use, the DNA solution was boiled for 15 min and kept on ice.
Fixer Solution	Mix 150 mL methanol + 36 mL acetic acid + 150 µL 37% formaldehyde and complete to 300 mL with distilled water. This solution can be used several times.
Pretreatment Solution	Dissolve 0.08 g sodium thiosulphate ($Na_2S_2O_3.5H_2O$) in 400 mL distilled water by mixing with a glass rod. Take 8 mL and set aside for further use in developing solution preparation.
Silver Nitrate Solution	Dissolve 0.8 g silver nitrate in 400 mL distilled water and add 300 µL 37% formaldehyde
Developing Solution	Dissolve 9 g potassium carbonate in 400 mL distilled water. Add 8 mL from pretreatment solution and 300 µL 37% formaldehyde.
Stop Solution	Mix 200 mL methanol + 48 mL acetic acid and complete to 400 mL with distilled water
Antifoam	10 % (v/v) antifoam solution, prepared with dH_2O. Can be autoclaved once.
Base for Bioreactor	25 % NH_3OH (Sigma). No need to sterilize.

APPENDIX D

DNA SEQUENCES AND PLASMIDS

Sequence of pPICZaA (3593 bp)

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacat
aagtgccaaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcct
caacacccacttttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttct
attaggctactaacaccatgactttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttc
cgaatgcaacaagctccgcattacacccgaacatcactccagatgagggctttctgagtgtggggtcaaatagttt
catgttccccaaatggcccaaaactgacagtttaaacgctgtcttggaacctaatatgacaaaagcgtgatctcat
ccaagatgaactaagtttggttcgttgaaatgctaacggccagttggtcaaaaagaaacttccaaaagtcggcata
ccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatgcttagcgcagtctctctatcg
cttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgctttttggatgattatgcattgtctc
cacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactgttct
aacccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattatta
gcttactttcataattgcgactggttccaattgacaagcttttgatttaacgacttttaacgacaacttgagaag
atcaaaaaacaactaattattcgaaacgatgagatttccttcaatttttactgctgttttattcgcagcatcctcc
gcattagctgctccagtcaacactacaacagaagatgaaacggcacaaattccggctgaagctgtcatcggttact
cagatttagaaggggatttcgatgttgctgttttgccattttccaacagcacaaataacgggttattgtttataaa
tactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaagagaggctgaagctgaattcacg
tggcccagccggccgtctcggatcggtacctcgagccgcggcggccgccagctttctagaacaaaaactcatctca
gaagaggatctgaatagcgccgtcgaccatcatcatcatcatcattgagtttgtagccttagacatgactgttcct
cagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtcagaatgccatttgcc
tgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggatttttttgtcattttgt
ttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggtaggggtttgggaaaa
tcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaagtgagaccttcgttt
gtgcggatcccccacacaccatagcttcaaaatgtttctactccttttttactcttccagattttctcggactccg
cgcatcgccgtaccacttcaaaacacccaagcacagcatactaaattttccctctttcttcctctagggtgtcgtt
aattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttcttttttcttcgtcgaaaaaggcaataa
aaatttttatcacgtttctttttcttgaaatttttttttttagtttttttctctttcagtgacctccattgatatt
taagttaataaacggtcttcaatttctcaagtttcagtttcatttttcttgttctattacaacttttttttacttct
tgttcattagaaagaaagcatagcaatctaatctaaggggcggtgttgacaattaatcatcggcatagtatatcgg
catagtataatacgacaaggtgaggaactaaaccatggccaagttgaccagtgccgttccggtgctcaccgcgcgc
gacgtcgccggagcggtcgagttctggaccgaccggctcgggttctcccgggacttcgtggaggacgacttcgccg
gtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggaccaggtggtgccggacaacaccctggcctg
ggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaacttccgggacgcctcc
gggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgccctgcgcgacccggccggcaactgcg
tgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcccaggcctcggagatccgtccc
ccttttcctttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctccccccacatccgctctaac
cgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagtattaagaacg
ttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgc
ttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagcaaaaggccagcaaaa
ggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgcccccctgacgagcatcacaaaaat
cgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcg
tgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttc
tcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccc
gttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccac
tggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcc
taactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagtt
ggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagcagcagattacgcgca
gaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgtta
agggattttggtcatgagatc

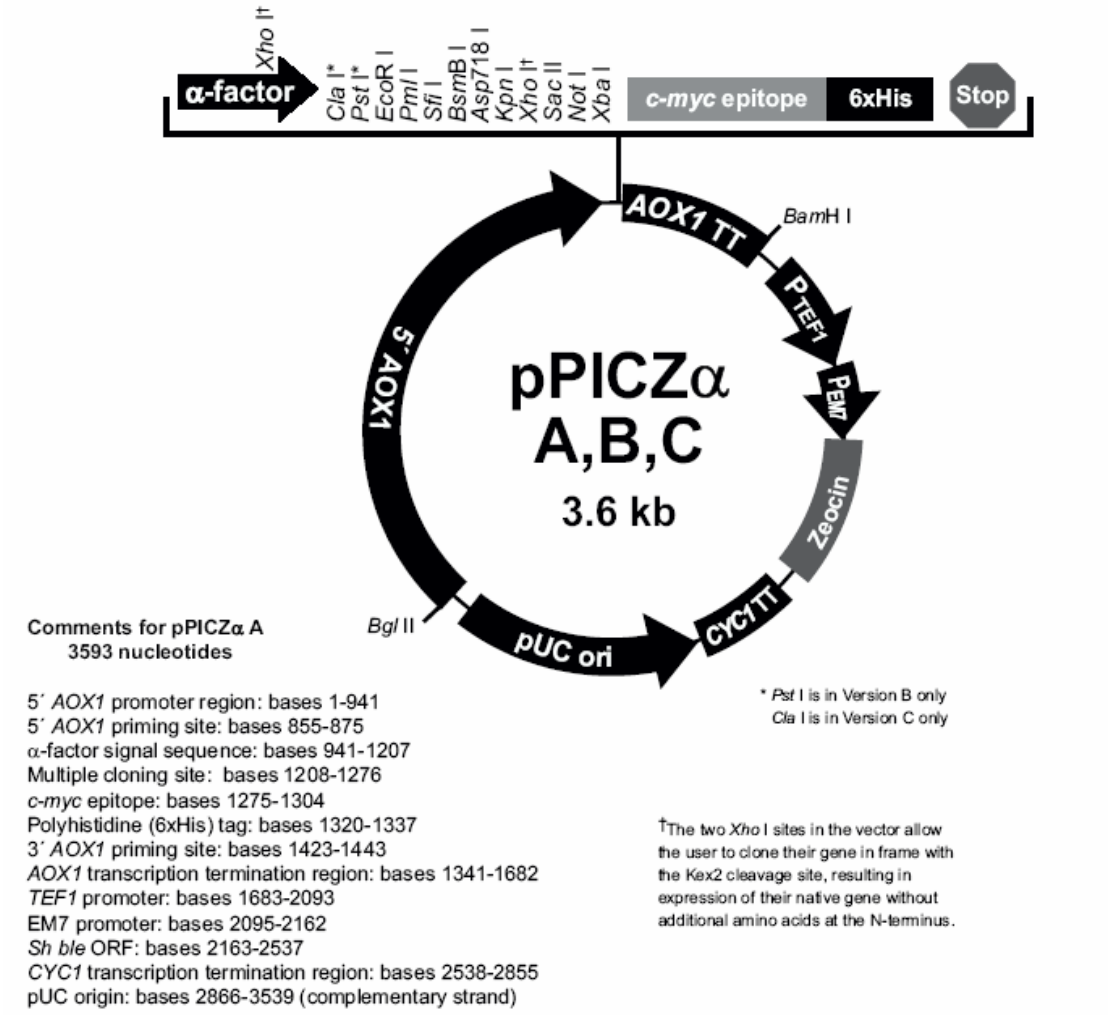
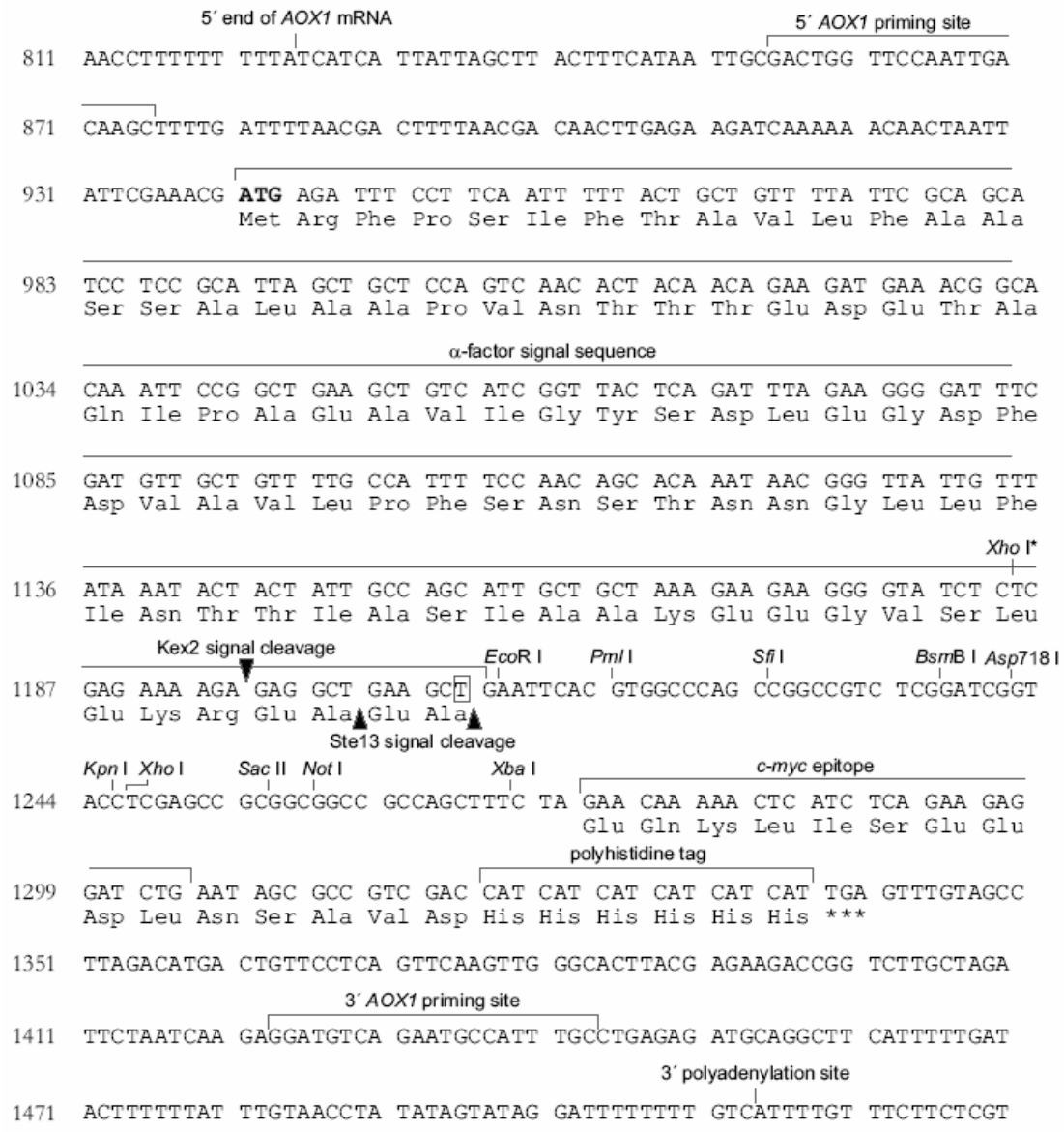


Figure D1. Schematic representation of pPICZαA vector as supplied from Invitrogen.

Multiple Cloning Site of pPICZαA



Sequence of pPICZαA-EPO (4064 bp)

agatctaacatccaaagacgaaaggttgaatgaaacctttttgccatccgacatccacaggtccattctcacacat
aagtgccaaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctcttctcct
caacacccacttttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattccttct
attaggctactaacaccatgactttattagcctgtctatcctggcccccctggcgaggttcatgtttgtttatttc
cgaatgcaacaagctccgcattacacccgaacatcactccagatgagggctttctgagtgtggggtcaaatagttt
catgttccccaaatggcccaaaactgacagtttaaacgctgtcttggaacctaatatgacaaaagcgtgatctcat
ccaagatgaactaagtttggttcgttgaaatgctaacggccagttggtcaaaaagaaacttccaaaagtcggcata
ccgtttgtcttgtttggtattgattgacgaatgctcaaaaataatctcattaatgcttagcgcagtctctctatcg
cttctgaaccccggtgcacctgtgccgaaacgcaaatggggaaacacccgctttttggatgattatgcattgtctc
cacattgtatgcttccaagattctggtgggaatactgctgatagcctaacgttcatgatcaaaatttaactgttct
aacccctacttgacagcaatatataaacagaaggaagctgccctgtcttaaacctttttttttatcatcattatta
gcttactttcataattgcgactggttccaattgacaagcttttgatttaacgacttttaacgacaacttgagaag
atcaaaaaacaactaattattcgaaacgatgagatttccttcaatttttactgctgttttattcgcagcatcctcc
gcattagctgctccagtcaacactacaacagaagatgaaacggcacaaattccggctgaagctgtcatcggttact
cagatttagaaggggatttcgatgttgctgttttgccattttccaacagcacaaataacgggttattgtttataaa
tactactattgccagcattgctgctaaagaagaaggggtatctctcgagaaaagagaggctgaagctgaattccac
catcaccatcaccatattgaagggagagccccaccacgcctcatctgtgacagccgagtcctggagaggtacctct
tggaggccaaggaggccgagaatatcacgacgggctgtgctgaacactgcagcttgaatgagaatatcactgtccc
agacaccaaagttaatttctatgcctggaagaggatggaggtcgggcagcaggccgtagaagtctggcagggcctg
gccctgctgtcggaagctgtcctgcggggccaggccctgttggtcaactcttcccagccgtgggagcccctgcagc
tgcatgtggataaagccgtcagtggccttcgcagcctcaccactctgcttcgggctctgggagcccaggaggaagc
catctcccctccagatgcggcctcagctgctccactccgaacaatcactgctgacactttccgcaaactcttccga
gtctactccaatttcctccggggaaagctgaagctgtacacaggggaggcctgcaggacaggggactaatctagaa
caaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcattgagtttgtagccttа
gacatgactgttcctcagttcaagttgggcacttacgagaagaccggtcttgctagattctaatcaagaggatgtc
agaatgccatttgcctgagagatgcaggcttcatttttgatacttttttatttgtaacctatatagtataggattt
tttttgtcattttgtttcttctcgtacgagcttgctcctgatcagcctatctcgcagctgatgaatatcttgtggt
aggggtttgggaaaatcattcgagtttgatgtttttcttggtatttcccactcctcttcagagtacagaagattaa
gtgagaccttcgtttgtgcggatccccacacaccatagcttcaaaatgtttctactccttttttactcttccaga
ttttctcggactccgcgcatcgccgtaccacttcaaaacacccaagcacagcatactaaattttccctctttcttc
ctctagggtgtcgttaattacccgtactaaaggtttggaaaagaaaaaagagaccgcctcgtttcttttttcttcgt
cgaaaaaggcaataaaaatttttatcacgtttctttttcttgaaatttttttttttagtttttttctctttcagtg
acctccattgatatttaagttaataaacggtcttcaatttctcaagtttcagtttcatttttcttgttctattaca
actttttttacttcttgttcattagaaagaaagcatagcaatctaatctaaggggcggtgttgacaattaatcatc
ggcatagtatatcggcatagtataatacgacaaggtgaggaactaaaccatggccaagttgaccagtgccgttccg
gtgctcaccgcgcgcgacgtcgccggagcggtcgagttctggaccgaccggctcgggttctcccgggacttcgtgg
aggacgacttcgccggtgtggtccgggacgacgtgaccctgttcatcagcgcggtccaggaccaggtggtgccgga
caacaccctggcctgggtgtgggtgcgcggcctggacgagctgtacgccgagtggtcggaggtcgtgtccacgaac
ttccgggacgcctccgggccggccatgaccgagatcggcgagcagccgtgggggcgggagttcgccctgcgcgacc
cggccggcaactgcgtgcacttcgtggccgaggagcaggactgacacgtccgacggcggcccacgggtcccaggcc
tcggagatccgtccccctttccctttgtcgatatcatgtaattagttatgtcacgcttacattcacgccctccccc
cacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatg
ttagtattaagaacgttatttatatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacatta
tactgaaaaccttgcttgagaaggttttgggacgctcgaaggctttaatttgcaagctggagaccaacatgtgagc
aaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccccctgac
gagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccc
ctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcggg
aagcgtggcgctttctcaatgctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggctgt
gtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagac
acgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagtt
cttgaagtggtggcctaactacggctacactagaaggacagtatttggtatctgcgctctgctgaagccagttacc
ttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgcaagc
agcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaa
cgaaaactcacgttaagggattttggtcatgagatc

Sequence of pENTR^{TM}221 (2546 bp without ORF)

ctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctc
gccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaagagcgcccaatacgcaaa
ccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaagc
gggcagtgagcgcaacgcaattaatacgcgtaccgctagccaggaagagtttgtagaaacgcaaaaag
gccatccgtcaggatggccttctgcttagtttgatgcctggcagtttatggcgggcgtcctgcccgcc
accctccgggccgttgcttcacaacgttcaaatccgctcccggcggatttgtcctactcaggagagcg
ttcaccgacaaacaacagataaaacgaaaggcccagtcttccgactgagcctttcgttttatttgatg
cctggcagttccctactctcgcgttaacgctagcatggatgttttcccagtcacgacgttgtaaaacg
acggccagtcttaagctcgggccccaaataatgattttattttgactgatagtgacctgttcgttgca
acaaattgatgagcaatgctttttataatgccaactttgtacaaaaaagcaggcaccacccagcttt
cttgtacaaagttggcattataagaaagcattgcttatcaatttgttgcaacgaacaggtcactatca
gtcaaaataaaatcattatttgccatccagctgatatcccctatagtgagtcgtattacatggtcata
gctgtttcctggcagctctggcccgtgtctcaaaatctctgatgttacattgcacaagataaaaatat
atcatcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgttatgagccatattc
aacgggaaacgtcgaggccgcgattaaattccaacatggatgctgatttatatgggtataaatgggct
cgcgataatgtcgggcaatcaggtgcgacaatctatcgcttgtatgggaagcccgatgcgccagagtt
gtttctgaaacatggcaaaggtagcgttgccaatgatgttacagatgagatggtcagactaaactggc
tgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactc
accactgcgatccccggaaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatat
tgttgatgcgctggcagtgttcctgcgccggttgcattcgattcctgtttgtaattgtccttttaaca
gcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttggttgatgcgagtgat
tttgatgacgagcgtaatggctggcctgttgaacaagtctggaaagaaatgcataaacttttgccatt
ctcaccggattcagtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaat
taataggttgtattgatgttggacgagtcggaatcgcagaccgataccaggatcttgccatcctatgg
aactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaatatggtattgataatcc
tgatatgaataaattgcagtttcatttgatgctcgatgagtttttctaatcagaattggttaattggt
tgtaacactggcagagcattacgctgacttgacgggacggcgcaagctcatgaccaaaatcccttaac
gtgagttacgcgtcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcc
ttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgc
cggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatact
gtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgc
tctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaa
gacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttg
gagcgaacgacctacaccgaactgagatacctacagcgtgagcattgagaaagcgccacgcttcccga
agggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttc
cagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgattt
ttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcct
ggcctttgctggccttttgctcacatgtt

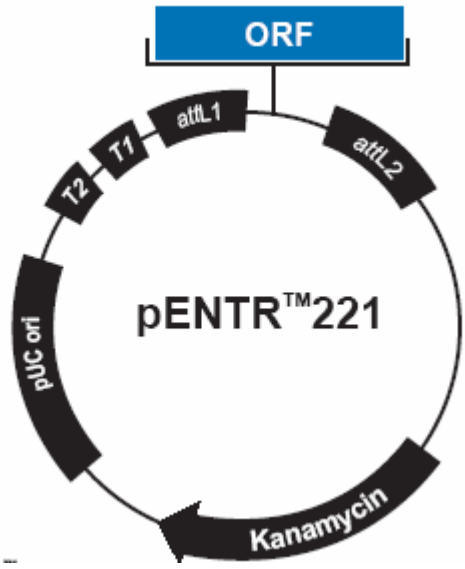


Figure D2. Schematic representation of pENTR 221 as supplied from Invitrogen

APPENDIX E

GROWTH MEDIA

LB (Luria-Bertani)

10 g L^{-1} NaCl
10 g L^{-1} Tryptone
5 g L^{-1} Yeast extract

After adjusting pH to 7.0 with NaOH, the final volume is adjusted with distilled water. 20 g L^{-1} agar is added if solid medium is required. The medium is autoclaved and stored at room temperature. In media referred to as LB+Kanamycin, Kanamycin was added after sterilization when the medium has cooled to about 55°C, at a final concentration of 50 µg mL^{-1} from 50 mg mL^{-1} stock.

LSLB (Low Salt Luria-Bertani)

5 g L^{-1} NaCl

10 g L^{-1} Tryptone

5 g L^{-1} Yeast extract

After adjusting pH to 7.0 with NaOH, the final volume is adjusted with distilled water. 20 g L^{-1} agar is added if solid medium is required. The medium is autoclaved and stored at room temperature. In media referred to as LSLB+Zeocin, Zeocin was added after sterilisation when the medium has cooled to about 55°C, at a final concentration of 25 µg mL^{-1} from 100 mg mL^{-1} stock.

BMGY or BMMY

10 g L^{-1} Yeast extract
20 g L^{-1} Peptone
0.1 M Potassium phosphate buffer pH 6.0
13.4 g L^{-1} YNB
4×10^{-5} g L^{-1} Biotin
10 g L^{-1} Glycerol (for BGMY) or 5 mL L^{-1} Methanol (for BMMY)

5g yeast extract and 10g peptone was dissolved in dH_20 and the volume adjusted to 375 mL. The solution was autoclaved and allowed to cool to room temperature. 50 mL of 1M potassium phosphate buffer stock pH 6.0, 25 mL of 20X YNB stock and 1mL of 500X Biotin (filter sterilized) was added. Thereafter, either 50 mL of 10% Glycerol (autoclaved) for BMGY media (precultivation media) or 50 mL of 10% Methanol for BMMY media (production media) was added.

YPD

10 g L^{-1} Yeast extract
20 g L^{-1} Peptone
20 g L^{-1} Glucose

Glucose is autoclaved separately as a 40% solution. 20 g L^{-1} agar is added if solid medium is required. The media are autoclaved and stored at room temperature. 50 mL of 40% glucose solution is added to 950 mL yeast extract and peptone solution when required. In media referred to as YPD+Zeocin, Zeocin was added after sterilization when the medium has cooled to about 55°C, at a final concentration of 100 µg mL^{-1} from 100 mg mL^{-1} stock.

SOC

20 g L^{-1} Bacto-tryptone
5 g L^{-1} Yeast extract
0.5 g L^{-1} NaCl
0.19 g L^{-1} KCl
0.02 M Glucose

The pH was adjusted to 7.0 with 10N NaOH and the final volume adjusted with dH_20. 1M Glucose solution was autoclaved separately and added after sterilization. The medium was aliquoted into microcentrifuge tubes aseptically and stored at room temperature.

Macronutrient Solutions

Table E.1 Composition of the various macronutrient solutions

	kg m^{-3} in medium				kg m^{-3} in 5X Stock Solution			
	BSM	mBSM	m2BSM	m3BSM	BSM	mBSM	m2BSM	m3BSM
$MgSO_4 7H_2O$	14.9	3.73	4.7	1.18	74.5	18.65	23.5	5.9
$CaSO_4$	0.93	0.23	-	-	4.65	1.15	-	-
$CaCl_2.2H_2O$	-	-	0.36	0.11	-	-	1.8	0.55
Reference	*a*	*b*	*c*	*d*				

a: Sibirnyi et al., 1987
b: Brady et al, 2001
c: Thorpe et al., 1999
d: Jungo et al, 2006

Trace Element Solutions

Table E.2 Composition of the various trace element solutions

	Conc. in medium in g m^3			Conc. in stock soln. in kg m^3		
Component	PTM4	PTM1	PTMJ	PTM4	PTM1	PTMJ
$CuSO_4 5H_2O$	8.7	26.1	8	2	6	4
NaI	0.3	0.3	-	0.08	0.08	-
KI	-	-	0.24	-	-	0.12
$MnSO_4.H_2O$	13.0	13.0	28	3	3	14
$Na_2MoO_4.2H_2O$	0.9	0.9	5.2	0.2	0.2	2.6
H_3BO_3	0.1	0.1	8	0.02	0.02	4
$ZnCl_2$	30.4	87.0	-	7	20	-
$ZnSO_4.7H_2O$	-	-	44	-	-	22
$FeSO_4.7H_2O$	95.7	282.6	-	22	65	-
$FeCl_3.6H_2O$	-	-	75	-	-	37.5
$CoCl_2$	2.2	2.2	-	0.5	0.5	-
$CoCl_2.6H_2O$	-	-	8	-	-	4
H_2SO_4 (ml)	4.3	21.7	-	1	5	-
Reference	*e*	*f*	*g*			

e: McGrew et al, 1997
f: Sibirnyi et al., 1987
g: Jungo et al, 2006

Basal Salt Medium (BSM)

40 g L^{-1}	Glycerol
26.7 mL L^{-1}	85% (w/w) H_3PO_4
0.93 g L^{-1}	$CaSO_4$
18.2 g L^{-1}	K_2SO_4
14.9 g L^{-1}	$MgSO_4 \cdot 7H_2O$
4.13 g L^{-1}	KOH

Sterilized by autoclave. This is a clear solution with pH_i=2.0.

PTM1 Salt Solution

6.0 g L^{-1} $CuSO_4$
0.08 g L^{-1} KI
3.0 g L^{-1} $MnSO_4$
0.2 g L^{-1} Na_2MoO_4
0.02 g L^{-1} H_3BO_3
0.5 g L^{-1} $CoCl_2$
20.0 g L^{-1} $ZnCl_2$
65.0 g L^{-1} $FeSO_4 \cdot 7H_2O$
0.2 g L^{-1} Biotin
5 mL L^{-1} H_2SO_4, 98% (w/w)

Filter sterilized. Results in turquoise color clear solution. Keep at +4℃. Discard when the color turns yellowish-green.

Sorbitol Stock Solution for Bioreactor

450 g L^{-1} Sorbitol

Dissolve 45 g sorbitol in dH2O and dilute till 100 mL (needs about 70 mL dH2O). Sterilize by autoclave. Add to 800 mL bioreactor medium.

Methanol Feed

Pure methanol and 12 mL L^{-1} sterile PTM1. No need to sterilize

Glycerol Feed

100 mL pure glycerol diluted to 200 mL with dH_2O and sterilized by autoclave.

APPENDIX F

MOLECULAR WEIGHT MARKERS

A. **B.** **C.**

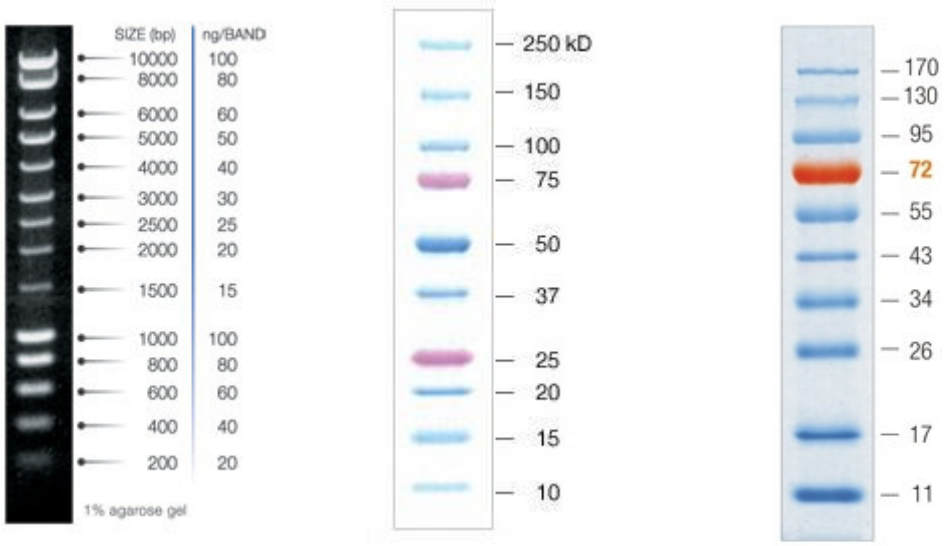


Figure F.1 **A.** DNA Ladder, Hyperladder I (BioLine) ; **B.** Precision Plus Dual Color Prestained Protein MW Marker (BioRad) **C.** PageRuler™ Prestained Protein Ladder (Fermentas)

APPENDIX G

AMINO ACID CODONS, ABBREVIATIONS AND CODON USAGE BY *P. pastoris*

Table G.1 Amino Acid Codons and Abbreviations

Amino acid	One-letter code	Three-letter-code	Codons
Alanine	A	Ala	GCU, GCC, GCA, GCG
Cysteine	C	Cys	UGU, UGC
Aspartic Acid	D	Asp	GAU, GAC
Glutamic Acid	E	Glu	GAA, GAG
Phenylalanine	F	Phe	UUU, UUC
Glycine	G	Gly	GGU,GGC,GGA,GGG
Histidine	H	His	CAU, CAC
Isoleucine	I	Ile	AUU, AUC, AUA
Lysine	K	Lys	AAA, AAG
Leucine	L	Leu	CUU, CUC, CUA, CUG
Methionine	M	Met	AUG (start codon)
Asparagine	N	Asn	AAU, AAC
Proline	P	Pro	CCU, CCC, CCG,CCA,
Glutamine	Q	Gln	CAA, CAG
Arginine	R	Arg	CGU, CGC,CGA, CGG
Serine	S	Ser	AGU, AGC
Threonine	T	Thr	ACU, ACC, ACA, ACG
Valine	V	Val	GUU, GUC, GUA, GUG
Tryptophan	W	Trp	UGG
Tyrosine	Y	Tyr	UAU, UAC
Stop Codons			UGA, UAA, UAG

Table G.2 Codon usage by *P. pastoris* (http://www.kazusa.or.jp/codon/, Last accessed: June 2008).

UUU 23.5	UCU 23.9	UAU 14.8	UGU 8.0
UUC 20.2	UCC 16.6	UAC 18.7	UGC 4.4
UUA 15.3	UCA 14.6	UAA 0.9	UGA 0.3
UUG 32.1	UCG 6.9	UAG 0.5	UGG 10.3
CUU 15.6	CCU 15.0	CAU 11.3	CGU 6.8
CUC 7.4	CCC 6.4	CAC 9.2	CGC 2.1
CUA 10.9	CCA 17.5	CAA 24.2	CGA 4.3
CUG 15.4	CCG 3.9	CAG 14.7	CGG 1.9
AUU 31.0	ACU 22.5	AAU 23.0	AGU 12.3
AUC 20.1	ACC 14.6	AAC 25.8	AGC 7.4
AUA 11.4	ACA 13.4	AAA 29.6	AGA 20.6
AUG 18.6	ACG 6.1	AAG 35.4	AGG 6.5
GUU 27.6	GCU 30.0	GAU 35.9	GGU 27.6
GUC 15.5	GCC 17.2	GAC 26.5	GGC 8.2
GUA 9.6	GCA 15.1	GAA 39.0	GGA 19.7
GUG 12.4	GCG 3.8	GAG 30.4	GGG 5.8

[triplet] [frequency: per thousand]

APPENDIX H

PROPERTIES OF DESIGNED PRIMERS

Table H.1 Thermodynamic properties of the designed primers.

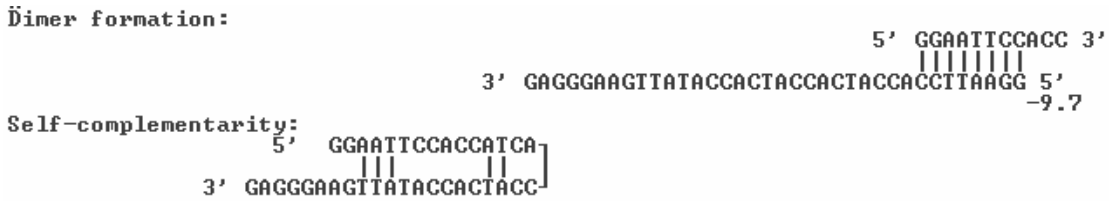
Name of Primer	Length (bp)	Gc (%)	T_m (°C)	ΔG (kcal/mol)	ΔH (kcal/mol)
EPO-F3-1	36	58	90.5	-68.5	-285.6
EPO-F3-2	36	47	85.4	-62.7	-271.7
EPO-R3	29	59	79.7	-51.1	-224.9
Seq-1-F	21	48	64.1	-32.9	-153.2
Seq-1-R	21	48	67.3	-34.8	-157.4

The self complimentary and dimer formation affinities of the primers are given below:

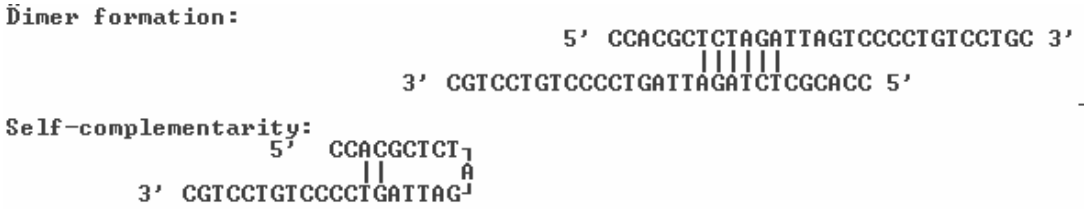
EPO-F3-1

```
Dimer formation:
                                                    5' CACCATATTGAAGGG 3'
                                                       | |||| |
                                  3' CTACTCCGCACCACCCCGAGAGGGAAGTTATACCAC 5'
                                                                      1.1
Self-complementarity:
                  5'  CACCATATTGAAGGGAG
                      | |  |||
                3' CTACTCCGCACCACCCCGA
```

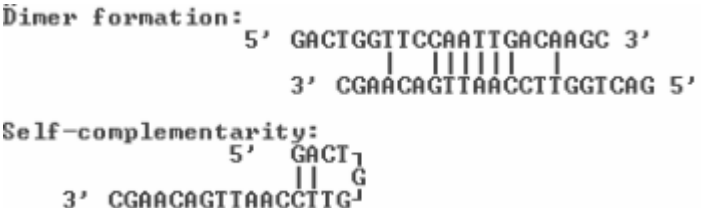
EPO-F3-2



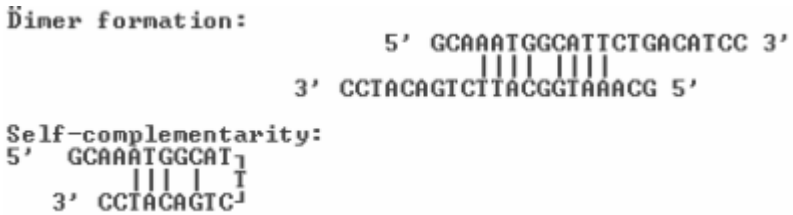
EPO-R3



Seq-1-F



Seq-1-R



APPENDIX I

PROTOCOL FOR SOUTHERN BLOTTING

1. Preparation and Electrophoresis of Genomic DNA

1) Isolate the yeast genomic DNA (Section 3.6.11) from selected colonies.

2) Digest the genomic DNA with a restriction endonuclease that is a single cutter of the vector (*Bam*HI), at 37°C overnight (Section 3.6.6).

3) To purify the DNA add 300 µL EtOH, vortex and leave on ice for 20 min. Centrifuge at maximum speed for 4 min, pour off the EtOH and incubate in 70°C water-bath for 10min to evaporate EtOH. Dissolve the pellet in 10 µL dH_2O and add 2µL loading dye to each sample.

4) Pour 200 mL, 0.7% agarose gel on 15x15 cm^2 electrophoresis tray.

5) Load the samples and run the gel at 60 V for 5h.

6) After the run is complete, take a picture of the gel with a ruler next to it.

2. Vacuum Blotting - VacuGene XL vacuum blotting system (Amersham)

1) Set-up the vacuum blotting unit.

2) Connect the pump and the liquid trap.

3) Cut the membrane (Hybond^{TM} N+, Amersham) as 15 x15 cm^2 and wet first in water then equilibrate in the Transfer buffer and cut the left hand corner for orientation.

4) Place the membrane on the porous screen of the blotting unit and place the plastic mask on the membrane in a way that it overlaps on each side of the membrane by approximately 5 mm.

5) Fit the top frame of the system and lock the four clamps.

6) Place the gel on the membrane, starting with one of the gel edges, gradually slide the gel from the support plate onto the membrane to fill the window. Trapped air bubbles must be removed and cracks in gel must be sealed with melted agarose.

7) Immobilize the gel by switching on the vacuum pump.

8) Without waiting for the vacuum to reach the working level (50 mbar), pour Depurination Solution (Appendix C) onto the center of the gel. Leave the gel completely covered for 20 min and stabilize the vacuum at 50 mbar.
9) Tilt the blotting unit to remove the liquid and pipette the residual liquid.
10) Pour on the Denaturation Solution to cover the gel, leave for 20 min and remove completely as before.
11) Pour on the Neutralizing Solution to cover the gel, leave for 20 min and remove completely as before.
12) Pour on the Transfer Solution to cover the gel, leave for 1-3 h and remove completely as before. Ensure that the gel remains immersed during transfer.
13) Turn off the pump. Mark the wells. Remove the gel.
14) Remove the membrane and wash with trasfer solution to eliminate agarose.
15) Place the membrane between blotting papers and dry in 80°C oven for 2 h.

3. Preparation of labeled probe

1) Amplify the EPO gene by PCR (section 3.6.5) using the primers EPO-F3-1 and EPO-R3.
2) Run the PCR product on agarose gel and purify the band with size of 528 bp from the gel.
3) Dilute the DNA to a concentration of 10ng μL^{-1} with dH_2O.
4) Dilute 20 µL cross-linker solution with 80µL dH_2O .
5) Place 10 µL of the diluted DNA sample in a microcentrifuge tube and denature by heating for 5 min in a vigorously boiling water bath.
6) Immediately cool the DNA on ice for 5 min. Spin briefly in a microcentrifuge to collect the contents at the bottom of the tube.
7) Add 10 µL of reaction buffer to the cooled DNA. Mix thoroughly but gently.
8) Add 2 µL labeling reagent. Mix thoroughly but gently.
9) Add 10 µL of the cross-linker working solution. Mix thoroughly. Spin briefly in a microcentrifuge to collect the contents at the bottom of the tube.
10) Incubate the reaction for 30 min at 37°C.
11) The probe can be used immediately or kept on ice for up to 2 h. For long term storage, labeled probes may be stored in 50%(v/v) Glycerol at -15°C to -30°C for up to six months.

4. Hybridization

1) Pre-heat the required volume of prepared AlkPhos Direct hybridization buffer to 55°C. The volume of buffer should be equivalent to 0.25 mL/cm^2 of membrane

2) Place the blot (i.e. membrane) into the hybridization buffer and prehybridize for at least 15 min at 55°C in hybridization oven.
3) Add the labeled probe to the buffer used for the pre-hybridization step. Use 10 ng probe per mL of buffer.
4) Hybridize at 55°C overnight in a hybridization oven.

5. Post hybridization stringency washes

1) Preheat the primary wash buffer to 55°C. This is used in excess at a volume of 2–5 mL/cm^2 of membrane.
2) Carefully transfer the blots to this solution and wash for 10 min at 55°C, with gentle agitation.
3) Perform a further wash in fresh, primary wash buffer at 55°C for 10 min
4) Place the blots in a clean container and add an excess of secondary wash buffer. Wash, with gentle agitation, for 5 min at room temperature.
5) Perform a further wash in fresh, secondary wash buffer at room temperature for 5 min.

6. Signal generation and detection

1) Drain the excess secondary wash buffer from the blots (by touching the corner of the blot against the box used for washing the blots or other convenient clean surface) and place them (sample side up) on a clean, non-absorbent, flat surface. Do not allow the blots to dry out.
2) Pipette detection reagent on to the blots (30–40 µL/cm^2) and leave for 5 min. Drain off excess detection reagent by touching the corner of the blot on to the non-absorbent surface.
3) Wrap the blots in clear plastic wrap. Place the blots DNA side up, in the film cassette.
4) Switch off the lights and place a sheet of autoradiography film on top of the blots. Close the cassette and expose for 1 hour at room temperature. The DNA side of the filter (wrapped in plastic wrap) must be placed next to the film for maximum sensitivity.
5) Remove the film and develop. If required, expose a second film for an appropriate length of time. For an initial experiment try a 1 hour exposure. The signal lasts for several days reaching a peak a few h after addition of the detection reagents. Subsequent exposures can be made with suitably adjusted exposure times to get optimum signal-to-noise ratio.

APPENDIX J

PROTOCOL FOR HIS-TAG PURIFICATION

1. Thoroughly resuspend the BD TALON Resin.
2. Immediately transfer the required amount of resin suspension to a sterile tube that will accommodate 10–20 times the resin bed volume.
3. Centrifuge at 700xg for 2 min to pellet the resin.
4. Remove and discard the supernatant.
5. Add 10 bed volumes of 1X Equilibration/Wash Buffer and mix briefly to pre-equilibrate the resin.
6. Recentrifuge at 700xg for 2 min to pellet the resin. Discard the supernatant.
7. Repeat Steps 5 and 6.
8. Add the clarified sample to the resin.
9. Gently agitate at room temperature for 20 min on a platform shaker to allow the polyhistidine-tagged protein to bind the resin.
10. Centrifuge at 700xg for 5 min.
11. Carefully remove as much supernatant as possible without disturbing the resin pellet.
12. Wash the resin by adding 10–20 bed volumes of 1X Equilibration/Wash Buffer. Gently agitate the suspension at room temperature for 10 min on a platform shaker to promote thorough washing.
13. Centrifuge at 700xg for 5 min.
14. Remove and discard the supernatant.
15. Repeat Steps 12–14.
16. Add one bed volume of the 1X Equilibration/Wash Buffer to the resin, and resuspend by vortexing.
17. Transfer the resin to a 2-mL gravity-flow column with an end-cap in place, and allow the resin to settle out of suspension.
18. Remove the end-cap, and allow the buffer to drain until it reaches the top of the resin bed, making sure no air bubbles are trapped in the resin bed.

19. Wash column once with 5 bed volumes of 1X Equilibration/Wash Buffer.

20. [Optional]: If necessary, repeat Step 19 under more stringent conditions using 5–10 mM imidazole in 1X Equilibration/Wash Buffer.

21. Elute the polyhistidine-tagged protein by adding 5 bed volumes of Elution Buffer to the column. Collect the eluate in 500-µL fractions.

Note: Under most conditions, the majority of the polyhistidine-tagged protein will be recovered in the first two bed volumes.

22. Use spectrophotometric and SDS/PAGE analyses to determine which fraction(s) contain(s) the majority of the polyhistidine-tagged protein.

23. Clean the resins with MES buffer to remove imidazole and then with dH_2O.

Store resin at 4°C in 20% non-buffered ethanol.

APPENDIX K

CALIBRATION OF PROTEIN CONCENTRATION

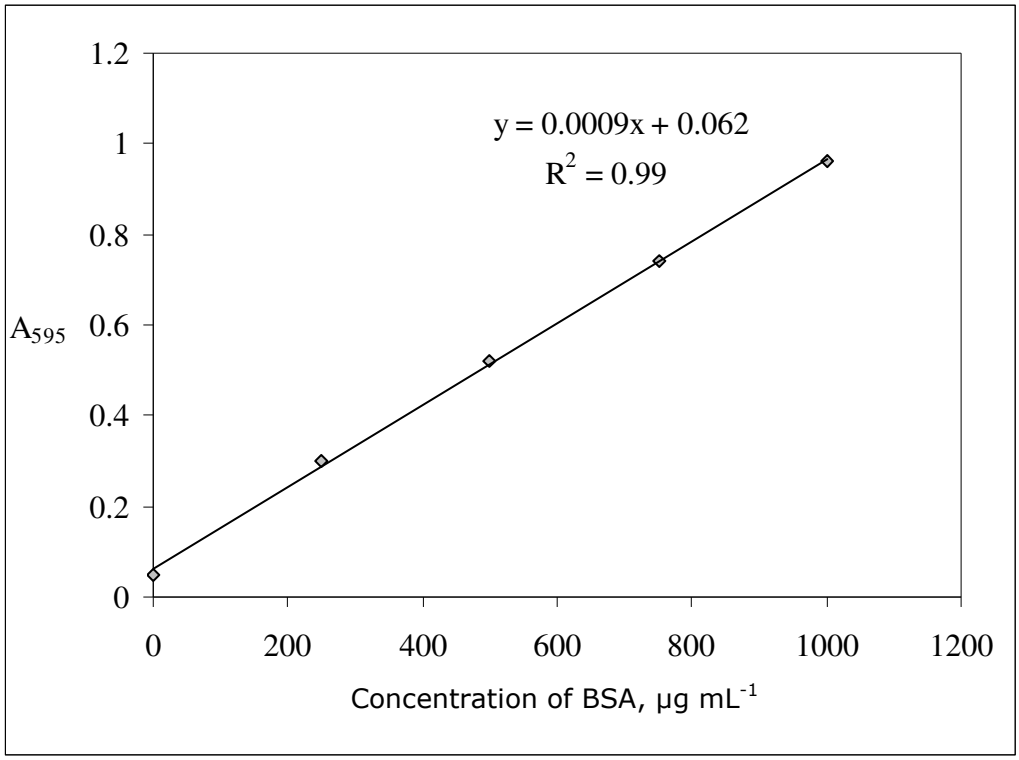


Figure K.1. Standard curve for Bradford Assay

$$Cp = \frac{A_{595} - 0.062}{9 \times 10^{-4}} \times \text{Dilution Factor}$$

APPENDIX L

SDS-PAGE AND WESTERN BLOTTING PROTOCOLS

Preparing the Gel

1. Set-up the gel cast.
2. Prepare the resolving and stacking gels without adding APS and TEMED.
3. Add TEMED and APS to the resolving gel, invert 2-3 times to mix gently, and quickly pipette into the gel cast to avoid polymerization but also avoid bubbles. Fill so that there will be 0.5 cm space below the comb-till the bottom of green line.
4. Pipette a thin layer of isopropanol quickly to smooth the surface.
5. Leave gel to polymerize for at least 45 min.
6. Pour away the isopropanol and wash 3 times with tap water and distilled water, shaking out excess water in btw washes and soak paper towel to dry.
7. Add TEMED and APS to the stacking gel, invert 2-3 times to mix gently, and quickly pipette into the gel cast till the top, avoid bubbles.
8. Insert the comb with an angle to avoid trapping bubbles.
9. Leave the gel for at least 20 min to polymerize.
10. Gel can be used straight away or stored in the fridge for up to 2 weeks, wrapped in tissue soaked in dH_2O and then wrapped in cling film.

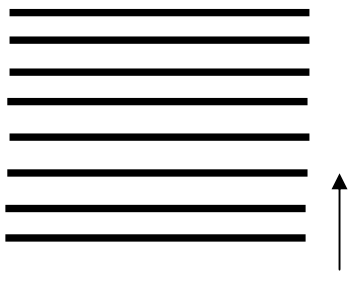
Gel electrophoresis

1. Add sample loading buffer to the samples (1:3) and heat at 95° C for 4 min, store on ice for 5 min, centrifuge and vortex.
2. Remove the comb slowly and clean the gel removing any excess with razor blade.
3. Wash the wells 3 times with tap water and distilled water, shaking out excess water in btw washes and soak paper towel to dry.
4. Assemble the gels onto the electrophoresis unit.

5. Fill the apparatus with 1x SDS-PAGE runnig buffer. Do not wet the electrodes. The space between the gels and the apparatus should be filled to the top, as well as the bottom of the apparatus to the point where the bottom of the gel is covered.
6. Load the samples and the prestained protein ladder into the wells with Pasteur pipette quickly and load the loading buffer into the unused wells as well, so that these lanes still visibly run. Make a note of each well and make individual gels identifiable from each other.
7. Close the lid by attaching the correct red/black power leads.
8. Run the gels at 20 mA for 1 mm gels and 25 mA for 1.5mm gels. This value needs to be multiplied by the number of gelds being run at the same time.

Blotting

1) Cut off stacking gel and nick top left-hand corner of resolving gel for orientation.
2) Transfer gel, while still attached to glass plate, into glass tray filled with at least 3 cm of 1x transfer buffer and peel off gently with a wet spatula, and keep for 15-20 min at RT to remove salts and SDS.
3) Cut membrane and 2 blotting papers to size, and cut one corner
4) Pre-wet PVDF or other hydrophobic membranes with methanol and transfer to the tray to equilibrate in transfer buffer for at least 10 min before blotting.
5) Put 2 sponges and 2 blotting papers into the tray to wet.
6) Open the cassette by releasing both latch tabs along the edge opposite the hinges. Place the opened cassette into the tray filled with transfer buffer and assemble the transfer stack so that the molecules will migrate toward the membrane. Make sure there are no bubbles in btw the gel and the membrane.



Cathode (-) – top – black
Sponge
Blotting paper
Gel (aligning ladder side with membrane's cut corner)
Membrane
Blotting paper
Sponge
Anode (+) – bottom – grey

7) Close the cassette and press lightly to lock the tabs.

8) The cassette must be oriented so that the hinges face up so the black side of each cassette faces the black cathode panel.

9) Throw a magnetic stirrer inside. Inspect the buffer level, should be btw min-max lines.

10) Close the lid, turn on the stirrer and run the gel in cold room, at 50 V for 3 h, regardless of the number of gels.

Blocking, Antibody Incubation and Detection Using ECL Western Blotting Kit

1) Remove the membrane from cassette, put in a box, protein side facing up. Wash the membrane 2-3 times with TBS-T. Tween prevents non-specific binding of the Ab and optimizes the hybridization conditions.

2) Immerse the membrane in TBS-T-Milk for 1 hr, RT, shaking.

3) Wash 3 times (shaking 15 min, 5min, 5 min with fresh changes) with large volumes of TBS-T, at RT.

4) Transfer the membrane to a lid of 96-well microtiter plate or similar low volume container.

5) Dilute the primary antibody in TBS-T as 1:1 000

6) Incubate the membrane in the diluted primary Ab for o/n at 4 $^{\circ}$C, on shaker platform.

7) Transfer the membrane into gel box and wash as in step 3.

8) Dilute the secondary antibody in TBS-T as 1:10 000

(Dilution factor should be determined empirically for each Ab as 1:1 000 - 1:10 000; more dilution will improve the linearity and increase the sensitivity).

9) Transfer the membrane back to the small container and incubate the membrane in the diluted secondary Ab for 1 hr at RT, on shaker platform.

10) Transfer the membrane into gel box and wash 3 times (10 min/wash) with fresh changes of large volumes of TBS-T buffer, at RT.

11) Mix the substrate components to prepare the chemiluminescent or immunocytochemical stain working solution. The final required volume is 0.125 mL/cm^2.

12) Place the membrane protein side-up on clear plastic wrap. Pipette the working substrate solution onto the membrane and wait for 5 min at room temperature.

15) Transfer the membrane to a new plastic wrap, protein side down. Wrap the membrane.

16) Take the picture (after developing the X-ray film in case of chemiluminescent substrate).

APPENDIX M

STAINING PROTOCOL FOR SDS-POLYACRYLAMIDE GELS WITH SILVER SALTS

	STEP	SOLUTION*	TIME OF TREATMENT	COMMENTS
1	Fixing	Fixer	≥ 1 hr	Overnight incubation is acceptable
2	Washing	50% Ethanol	3 x 20 min	Should be fresh
3	Pre-treatment	Pretreatment Solution	1 min	Should be fresh
4	Rinse	Distilled water	3 x 20 sec	Time should be exact
5	Impregnate	Silver Nitrate Solution	20 min	
6	Rinse	Distilled water	2 x 20 sec	Time should be exact
7	Developing	Developing Solution	~ 5 min	After a few minutes add some distilled water to proceed the reaction slowly. Time should be determined by observation of color development
8	Wash	Distilled water	2 x 2 min	
9	Stop	Stop Solution	≥ 10 min	The gels can be kept in this solution overnight

* Solutions are given in Appendix C.

APPENDIX N

CALIBRATION OF METHANOL, GLYCEROL AND SORBITOL CONCENTRATIONS

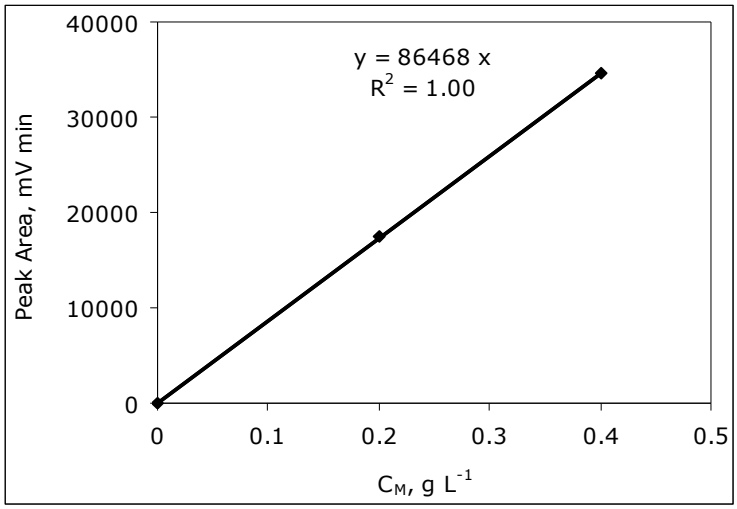


Figure N.1 Calibration curve obtained for methanol concentration; analysis was performed by HPLC.

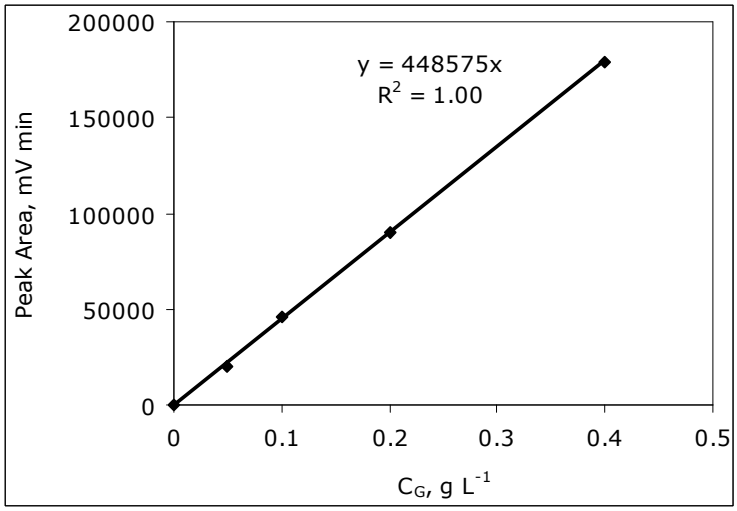


Figure N.2 Calibration curve obtained for glycerol concentration; analysis was performed by HPLC.

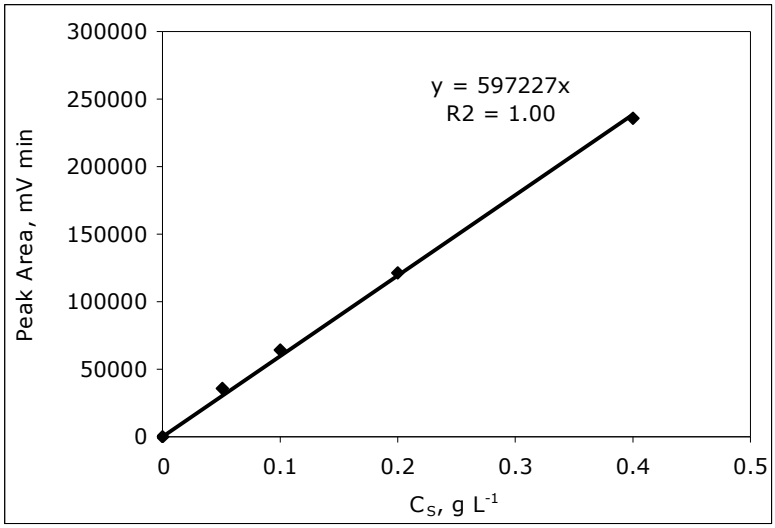


Figure N.3 Calibration curve obtained for sorbitol concentration; analysis was performed by HPLC.

APPENDIX O

METABOLIC REACTIONS FOR *P. pastoris*

Methanol Metabolism

1. MeOH → Form
2. Form → For + NADH
3. For → NADH + CO_2
4. Xyl5P + Form + ATP → ADP + 2 G3P

Sorbitol Metabolism

5. Sorb + ATP → ADP + F6P +NADH

Glycolysis and Gluconeogenesis Pathway

6. F6P → G6P
7. G6P → F6P
8. F6P + ATP → 2 G3P + ADP
9. 2 G3P → F6P + Pi
10. G3P + ADP + Pi → 3PG + ATP + NADH
11. 3PG + ATP + NADH→ G3P + ADP + Pi
12. 3PG → PEP
13. PEP → 3PG
14. PEP +ADP → Pyr + ATP
15. Pyr → AcCoAm + CO_2 + NADHm

Pentose Phosphate Pathway

16. G6P → R5P + 2 NADPH + CO_2
17. R5P → Xyl5P
18. Xyl5P → R5P
19. R5P → Rib5P
20. Rib5P → R5P
21. Xyl5P + Rib5P → S7P + G3P
22. S7P + G3P → Xyl5P + Rib5P
23. Xyl5P + E4P → F6P + G3P

24. F6P + G3P → Xyl5P + E4P
25. G3P +S7P → F6P + E4P
26. F6P + E4P → G3P +S7P

Branches from Glycolysis Pathway

27. Pyr → Acet + CO_2
28. Acet → Ac + NADPH
29. Acet + NADH → EtOH
30 Ac + 2 ATP → AcCoA + 2 ADP + 2Pi
31. AcCoA → AcCoAm
32. Pyr + NADH → Lac
33. Lac → NADH + Pyr

Anaplerotic Reactions

34. Mal → Pyr + CO_2 + NADPHm
35. Pyr + CO_2 + ATP → OA + ADP
36. OA + ATP → PEP + ADP + CO_2

TCA cycle

37. AcCoAm + OA → Cit
38. Cit → ICit
39. ICit → αKG +CO_2+ NADHm
40. αKG → SucCoA + CO_2 + NADHm
41. SucCoA + Pi + ADP → Suc + ATP
42. Suc + ATP → SucCoA + ADP + Pi
43. Suc → Fum + $FADH_2$
44. Fum → Mal
45. Mal → Fum
46. Mal → OA + NADHm
47. $NADH_m$ + OA → Mal

Biosynthesis of Serine Family Amino Acids

48. 3PG + Glu → Ser + αKG + NADH + Pi
49. Ser + THF → Gly + MetTHF
50. Ser + AcCoA + H_2S → Cys

Biosynthesis of Alanine Family Amino Acids

51. Pyr + Glu → Ala + αKG
52. 2 Pyr + NADPHm → Kval + CO_2
53. Kval + Glu → Val + αKG

54. Kval + AcCoAm + Glu → Leu + αKG + NADH + CO_2

Biosynthesis of Histidine

55. R5P + ATP → PRPP + AMP

56. PRPP + ATP + Gln → His + PRAIC + αKG + 2 PPi + 2 NADH + Pi

Biosynthesis of Aspartic Family Amino Acids

57. OA + Glu → Asp + αKG

58. Asp + Gln + ATP → Asn + Glu + AMP + PPi

59. Asp + ATP + 2 NADPH → Hser + ADP + Pi

60. HSer + ATP → Thr + ADP + Pi

61. Thr + NADPHm + Glu + Pyr → Ile + αKG + NH_4 +CO_2

62. AcCoA + HSer + H_2S + MTHF → Met + THF

Biosynthesis of Aromatic Family Amino Acids

63. 2 PEP + E4P + ATP + NADPH → Chor + ADP + 4 Pi

64. Chor + Glu → Phe + αKG + CO_2

65. Chor + Glu → Tyr + αKG + NADH + CO_2

66. Chor + Gln + PRPP + Ser → Trp + Glu + Pyr + G3P + CO_2 + PPi

Biosynthesis of Glutamic Family Amino Acids

67. αKG + NH_4 + NADPH → Glu

68. Glu + ATP + NH_4 → Gln + ADP + Pi

69. Glu + ATP + 2 NADPH → Pro + ADP + Pi

70. Gln + CO_2 + 2 ATP → CaP + Glu + 2 ADP + Pi

71. 2 Glu + AcCoAm + 4ATP + NADPHm + CaP + Asp → Arg + αKG + 4ADP + Fum + 5Pi

72. 2 Glu + AcCoA + 3 ATP + 2 NADPH → Lys + αKG + CO2 + 2NADH

Catabolism of Amino Acids

73. Ala + αKG → Pyr + Glu

74. Arg+ αKG + NADPH → Glu + 2 NH_4 + CO_2

75. Asn → Asp + NH_4

76. Asp + αKG + NADH → Glu + Mal

77. Cys → Pyr + NH_4+ H_2S

78. Gln → Glu + NH_4

79. Glu → NH_4 + NADH + αKG

80. Gly + MetTHF → Ser + THF

81. His +THF → Glu + F10THF + NH_4

82. Ile + αKG → Glu + $FADH_2$ + 2 NADH + CO_2 + AcCoA + SucCoA

83. Leu + αKG + ATP → Glu + $FADH_2$ + NADH + 2 AcCoA + ADP + Pi

84. Phe + NADH → Tyr

85. Pro → Glu + NADH + $FADH_2$

86. Ser → Pyr + NH_4

87. Thr → Gly + NADH + AcCoA

88. Trp + NADPH → 2 AcCoA + Ala + CO_2 + NH_4 + For + 2NADH + $FADH_2$

89. Tyr + αKG → Glu + Fum + 2 AcCoA + CO_2

90. Val + αKG + ATP → Glu + $FADH_2$ + 3 NADH + CO_2 + SucCoA

91. Lys + AcCoA + 2 αKG → 2 Glu + NADH + CO_2

Biosynthesis of Nucleotides

92. PRPP + 2 Gln + Asp + CO_2 + Gly + 4 ATP + F10THF → 2 Glu + PPi + 4 ADP + 4 Pi + THF + PRAIC + Fum

93. PRAIC + F10THF → IMP +THF

94. IMP + Gln + ATP → NADH + GMP + Glu + AMP + PPi

95. GMP + ATP → GDP + ADP

96. ATP + GDP → ADP + GTP

97. GTP + ADP → ATP + GDP

98. NADPH + ATP → dATP

99. NADPH + GDP + ATP → dGTP + ADP

100. IMP + GTP + Asp → GDP + Pi + Fum + AMP

101. AMP + ATP → 2 ADP

102. PRPP + Asp + CaP→ UMP + NADH + PPi + Pi + CO_2

103. UMP + ATP → UDP + ADP

104. UDP + ATP → ADP + UTP

105. UTP + NH_4 + ATP → CTP + ADP + Pi

106. CTP + ADP → CDP + ATP

107. CDP + ATP → CTP + ADP

108. CDP + ADP → CMP + ATP

109. ATP+ NADPH + CDP → dCTP + ADP

110. UDP + MetTHF + 3 ATP + NADPH → dTTP + DHF + 3 ADP + Ppi + Pi

Biosynthesis and Interconversion of One-carbon Units

111. DHF + NADPH → THF

112. Gly + THF → MetTHF + NH_4 + NADH + CO_2

113. MetTHF + NADH → MTHF

114. MetTHF → MeTHF + NADPH

115. MeTHF → F10THF

Oxidative Phosphorylation *(P/O=2)*

116. NADHm + 2 ADP + 2 Pi → 2 ATP

117. $FADH_2$ + ADP + Pi → ATP

Transport Reactions

118. CO_2 → exp

119. imp → CO_2

120. imp → NH_4

121. NH_4 → exp

122. 2ATP + 4 NADPH → AMP + ADP + H_2S + PPi + Pi

123. PPi → 2Pi

124. imp → Pi

125. Pi → exp

126. NADH → NADHm

Biosynthesis of Carbohydrate

127. ATP + G6P → ADP + 6 CARBH + 2 Pi

Biosynthesis of Lipids

128. ACCoA + ATP + CO2 → ADP + MaCoA + Pi

129. ACCoA + 7 MaCoA + 14 NADPH → 7 CO2 + PLM

130. NADPH + PLM + ATP → PLLM

131. ACCoA + 8 MaCoA + 16 NADPH → 8 CO2 + STE

132. NADPH + STE + ATP → OLE

133. 1.7 OLE + 4.4 PLLM + 1.4 PLM + STE → 8.5 FA

134. 2 FA + G3P → PA

135. CTP + PA + Ser → 2 Pi + CMP + PS

136. PS → CO2 + PE

137. PE + 3ATP + 3 Met → PC + 3 AcCoA + 3 H_2S + 3 HSer + 9 Pi

138. PA + CTP + G6P → CMP + PINS

Biomass Synthesis

139. 0.459 Ala + 0.161 Arg + 0.102 Asn + 0.297 Asp + 0.007 Cys + 0.105 Gln + 0.302 Glu + 0.290 Gly + 0.066 His + 0.193 Ile + 0.296 Leu + 0.286 Lys + 0.051 Met + 0.134 Phe + 0.165 Pro + 0.185 Ser + 0.191 Thr + 0.028 Trp + 0.102 Tyr + 0.265 Val + 0.051 AMP + 0.051 GMP + 0.067 UMP + 0.05 CMP + 0.0024 dATP + 0.0016 dGTP +

0.0016 dTTP + 0.0024 dCTP + 0.0101 OLE + 0.0081 PLM +0.0263 PLLM + 0.0061 STE + 0.0006 PA + 0.005 PINS + 0.002 PS + 0.005 PE + 0.006 PC + 2.5 CARBH + 23.917 ATP $\rightarrow$ Biomass + 23.917 ADP + 23.946 Pi

Erythropoietin Synthesis

140. 34 Ala + 15 Arg + 11 Asn + 10 Asp + 4 Cys + 8 Gln + 24 Glu + 14 Gly + 8 His + 12 Ile + 30 Leu + 9 Lys + 2 Met + 11 Phe + 12 Pro + 18 Ser + 19 Thr + 3 Trp + 5 Tyr + 17 Val + 1064 ATP $\rightarrow$ EPO + 1064 ADP + 1064 Pi

Maintenance Energy

141. ATP $\rightarrow$ ADP + Pi

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Çelik Akdur, Eda
Nationality: Turkish (TC)
Date and Place of Birth: 19 June 1980, Ankara
Marital Status: Married
Phone: +90 532 657 17 41
Email: eda_celik@yahoo.com

EDUCATION

Degree	Institution	Year of Graduation
MS	METU Chemical Engineering	2003
BS	METU Chemical Engineering	2001
High School	Mehmet Emin Resulzade Anatolian High School	1997

WORK EXPERIENCE

Year	Place	Enrollment
2001 - 2008	METU Department of Chemical Engineering	Teaching Asst.
2005 - 2006	Uni.of Manchester Faculty of Life Sciences, UK	Visiting Sci.
2000 July-Sept.	Procter&Gamble, Brussels, Belgium	Intern
1999 July-Aug.	ETI-Alumina, Seydisehir, Turkey	Intern

FOREIGN LANGUAGES

Advanced English, Beginner Level French

SCHOLARSHIPS AND AWARDS

Poster Award, 2nd place, in Biotech2008 Conference, Switzerland, 2008.

The Scientific and Technological Research Council of Turkey (TUBITAK), National PhD Scholarship, 2004-2008.

The Scientific and Technological Research Council of Turkey (TUBITAK), International Research Scholarship for PhD studies, 2005-2006.

Middle East Technical University, Department of Chemical Engineering, Hasan Orbey Research Award in Graduate Degree, 2003.

The Scientific and Technological Research Council of Turkey (TUBITAK), National MSc Scholarship, 2001-2003.

PUBLICATIONS

Journal Articles Covered By Science Citation Index

1. Çelik, E., Çalık, P., Oliver, S.G., Fed-batch methanol feeding strategy for recombinant protein production by *Pichia pastoris* in the presence of co-substrate sorbitol, *in preparation.*

2. Çelik, E., Çalık, P., Oliver, S.G., Metabolic flux analysis for recombinant protein production by *Pichia pastoris* using dual carbon sources: Effects of methanol feeding rate, *in preparation.*

3. Çelik, E., Çalık, P., Oliver, S.G., A structured kinetic model for recombinant protein production by Mut^+ strain of *Pichia pastoris, in preparation.*

4. Çelik, E., Özbay, N., Oktar, N., Çalık, P., 2008. Use of biodiesel byproduct crude glycerol as the carbon source for fermentation processes by recombinant *Pichia pastoris*. Industrial & Engineering Chemistry Research., 47, 2985-2990.

5. Özdamar, T.H., Şentürk, B., Yılmaz, Ö.D., Çalık, G., Çelik, E., Çalık, P., *in press*: 2008. Expression system for recombinant human growth hormone production from *Bacillus subtilis*, Biotechnology Progress

6. Çalık, P., Orman, M.A., Çelik, E., Halloran, S.M., Çalık, G., Ozdamar, T.H., 2008. Expression system for synthesis and purification of recombinant human growth hormone in *Pichia pastoris* and structural analysis by MALDI-ToF mass spectrometry. Biotechnology Progress, 24 (1), 221-226.

7. Çelik, E., Çalık, P., Halloran, S.M., Oliver, S.G., 2007. Production of recombinant human erythropoietin from *Pichia pastoris* and its structural analysis, Journal of Applied Microbiology, 103(6), 2084-2094.

8. Çelik, E., Çalık P., 2004. Bioprocess and oxygen transfer characteristics in beta-lactamase production by *Bacillus* species. Biotechnology Progress, 20(2), 491-499.

9. Çalık P., Çelik, E., Telli, I.E, Oktar, C., Özdemir, E., 2003. Protein based complex medium design for recombinant serine alkaline protease production by *Bacillus subtilis*. Enzyme and Microbial Technology, 33(7), 975-986.

Abstracts Published in the Proceedings of International Conferences

1. Çelik, E., Özbay, N., Oktar, N., Çalık, P., 22-23/5/2008. Recombinant protein production in *Pichia pastoris* using byproduct of biodiesel production. Biotech2008, Biopharmaceuticals: Why use yeasts?, Switzerland.
2. Çelik, E., Çalık, P., Halloran, M., Oliver, S.G. Expression and structural characterization of recombinant human erythropoietin from *Pichia pastoris*. Journal of Biotechnology, 131 (2): S153-S154 Suppl. S, SEP 2007. (13^{th} European Congress on Biotechnology, Spain).
3. Çelik, E., Wishart, J., Zhang, N., Calik, P., Oliver, S.G., 24-28/06/2006. Investigation of *Pichia pastoris* as a system for recombinant erythropoietin

production. 10th International Symposium on the Genetics of Industrial Microorganisms, Abstract Book, P-256. Czech Republic.

4. Çelik, E., Çalık, P., 8-11/09/2004. Effects of temperature on intracellular reaction rates in beta-lactamase production by *Bacillus licheniformis*. 5th European Conference on Biochemical Engineering Science, Abstract Book, P-90, Germany.

5. Çelik, E., Çalık, P., 24-29/08/2003. Effects of induction strategy and medium composition on beta-lactamase production by *Bacillu*s species. 11th European Congress on Biotechnology, Abstract Book, p234, Switzerland.

6. Çelik, E., Çalık, P., 13-15/05/2003. Production of β-lactamase by *Bacillus* species: effect of medium composition. 3rd Eastern Mediterranean Chemical Engineering Conference, Greece.