EFFECTS OF METABOLIC ENGINEERING OF ETHANOL UTILIZATION PATHWAY IN PICHIA PASTORIS ON RECOMBINANT PROTEIN PRODUCTION AND TRANSCRIPTION LEVELS IN CENTRAL METABOLISM

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

BY

İDİL DAYANKAÇ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
CHEMICAL ENGINEERING

SEPTEMBER 2021

Approval of the thesis:

EFFECTS OF METABOLIC ENGINEERING OF ETHANOL UTILIZATION PATHWAY IN *Pichia pastoris* ON RECOMBINANT PROTEIN PRODUCTION AND TRANSCRIPTION LEVELS IN CENTRAL METABOLISM

submitted by **İDİl DAYANKAÇ** in partial fulfillment of the requirements for the degree of **Master of Science** i**n Chemical Engineering, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Pınar Çalık
Head of the Department, **Chemical Engineering** _____

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

Prof. Dr. Tunçer H. Özdamar
Co-Supervisor, **Chemical Engineering, Ankara University** _____

Examining Committee Members:

Assoc. Prof. Dr. Eda Çelik Akdur
Chemical Engineering Dept., Hacettepe University _____

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

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

Assoc. Prof. Dr. Yeşim Soyer
Food Engineering Dept., METU _____

Assist. Prof. Dr. Harun Koku
Chemical Engineering Dept., METU _____

Date: 07.09.2021

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 : İdil Dayankaç

Signature :

ABSTRACT

EFFECTS OF METABOLIC ENGINEERING OF ETHANOL UTILIZATION PATHWAY IN *Pichia pastoris* ON RECOMBINANT PROTEIN PRODUCTION AND TRANSCRIPTION LEVELS IN CENTRAL METABOLISM

Dayankaç, İdil
Master of Science, Chemical Engineering
Supervisor : Prof. Dr. Pınar Çalık
Co-Supervisor: Prof. Dr. Tunçer H. Özdamar

September 2021, 210 pages

The aim of this MSc thesis is to increase the ethanol uptake rate and decrease by-product formation in the ethanol utilization (EUT) pathway and acquire a smoothly operating intracellular reaction network in the target microorganism. A metabolically engineered novel host, *Pichia pastoris*, was developed by overexpression of the two EUT pathway enzymes, alcohol dehydrogenase 2 (ADH2) and Acetyl-CoA synthetase (ACS1) that catalyze rate-limiting reactions, one by one and together, to understand how they affect the ethanol uptake rate, cell growth, and by-product formation. The metabolically engineered novel *Pichia pastoris* expression systems constructed were denoted by; i) *ADH2*-OE, ii) *ACS1*-OE, iii) *ADH2*-OE + *ACS1*-OE. The influence of *ADH2* overexpression on the growth of *P. pastoris* strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* was investigated in batch cultivations on 2% (v/v) ethanol, with two separate colonies carrying 3 and 9 *ADH2* gene copies, named as *ADH2*-OE-C3 and *ADH2*-OE-C9, respectively. At the end of the fermentation, compared to *P. pastoris* X-33 strain, 1.1-fold and 1.35-fold lower cell concentrations were obtained with *ADH2*-OE-C3 and *ADH2*-OE-C9,

respectively. At t = 36 h of the fermentation, compared to *P. pastoris* X-33, acetic acid concentrations were 1.46-fold and 3.97-fold higher in *ADH2*-OE-C3 and *ADH2*-OE-C9, respectively; while 1.18-fold and 2.28-fold higher acetaldehyde concentrations were obtained with *ADH2*-OE-C3 and *ADH2*-OE-C9, respectively. The effects of *ADH2* and *ACS1* overexpression on the growth of *P. pastoris* were investigated in batch cultivations on 1% (v/v) ethanol using metabolically engineered strains: i) *ADH2*-OE, ii) *ACS1*-OE, and iii) *ADH2*-OE and *ACS1*-OE. At the end of the fermentation, compared to *ADH2*-OE and *ADH2-OE+ACS1-OE,* 1.08-fold and 1.12-fold higher cell concentrations were obtained with *ACS1*-OE. Furthermore, specific ethanol uptake rate (q_{EtOH}) at t = 9 h for *ADH2*-OE and *ADH2*-OE + *ACS1*-OE strains were 0.52 and 0.35 g g^{-1} h^{-1}, respectively. Compared to *P. pastoris* X-33 (q_{EtOH} = 0.19 $gg^{-1}h^{-1}$), showing that *ADH2* overexpression significantly increased ethanol uptake rate. 1.49-fold and 1.78-fold decrease in excreted acetic acid concentrations compared to *P. pastoris* X-33 and *ADH2*-OE were obtained with *ACS1*-OE at t = 36 h of the fermentation. The effects of overexpression of *ADH2* and *ACS1* on transcription levels of central metabolism were also investigated. Three genes encoding crucial enzymes in ethanol utilization, *ADH2*, *ALD4*, and *ACS1*, were upregulated in all three metabolically engineered strains. Finally, to investigate the effect of *ACS1* overexpression on r-protein production, a novel host strain was constructed in which the gene of the reporter r-protein *mApple* was expressed under the control of $P_{ADH2\text{-}Cat8\text{-}L2}$, and *ACS1* was overexpressed under $P_{AOX1/Cat8\text{-}L3}$. *ACS1*-OE-C4 increased mApple production 1.2-fold at t= 24 h of the fermentation, and 1.32-fold at t=45 h of the fermentation.

Keywords: *Pichia pastoris*, Metabolic engineering, Ethanol utilization pathway, Transcription level analysis

ÖZ

Pichia pastoris'in ETANOL TÜKETİM YOLUNDA YAPILAN METABOLİK MÜHENDİSLİĞİNİN REKOMBİNANT PROTEİN ÜRETİMİ VE MERKEZİ METABOLİZMADAKİ TRANSKRİPSİYON DÜZEYLERİNE ETKİSİ

Dayankaç, İdil
Yüksek Lisans, Kimya Mühendisliği
Tez Yöneticisi: Prof. Dr. Pınar Çalık
Ortak Tez Yöneticisi: Prof. Dr. Tunçer H. Özdamar

Eylül 2021, 210 sayfa

Bu yüksek lisans tezinde, etanol tüketim hızının artırılması ve etanol yolizi üzerindeki yan-ürünlerin üretiminin azaltılması ile iyi işleyen hücre içi reaksiyon ağının elde edilmesi hedeflenmiştir. Metabolik mühendislik ile etanol yolizi üzerindeki hız kısıtlayıcı reaksiyonları katalizleyen iki enzimin, alkol dehidrogenaz (ADH2) ve Asetil Koenzim A sentaz (ACS1), ayrı ayrı ve beraber aşırı ekspresyonu ile özgün konak *Pichia pastoris* geliştirilmiştir. Metabolik mühendislik ile geliştirilen yeni *Pichia pastoris* ekspresyon sistemlerinin; i) *ADH2*-OE, ii) *ACS1*-OE, iii) *ADH2*-OE ve *ACS1*-OE hücre çoğalması, etanol tüketim hızı ve yan-ürün oluşumuna etkileri araştırılmıştır. *ADH2* aşırı ekspresyonunun, $P_{ADH2\text{-}wt::\,ADH2}$ kaseti ile oluşturulan *P. pastoris* suşlarının hücre çoğalması üzerine etkisi, üç *ADH2* gen kopyası ve dokuz *ADH2* gen kopyası taşıyan kolonileri ile %2 (h/h) başlangıç etanol derişiminde araştırılmıştır. Doğal *P. pastoris* X-33 ile karşılaştırıldığında, üç ve dokuz *ADH2* gen kopyası taşıyan hücreler ile sırasıyla 1.11 kat ve 1.35 kat daha düşük hücre derişimi elde edilmiştir. Fermentasyonun t = 36 st'te, *P. pastoris* X-33 ile üretimle kıyaslandığında, üç ve dokuz *ADH2* gen kopyası taşıyan hücreler ile yan-

ürün asetik asit derişimleri sırasıyla 1.46 kat ve 3.97 kat, yan-ürün asetaldehit derişimleri de sırasıyla 1.18 kat ve 2.28 kat artmıştır. *ADH* ve *ACS1* aşırı ekspresyonlarının, i)*ADH2*-OE, ii) *ACS1*-*OE* ve iii) ADH2-*OE* ve *ACS1*-OE ile oluşturulan *P. pastoris* suşlarının çoğalması üzerindeki etkisi, %1 (h/h) başlangıç etanol derişiminde araştırılmıştır. Fermentasyon sonunda *ADH2*-OE ve *ADH2*-OE+*ACS1*-OE'ye göre *ACS1*-OE ile 1.08 kat ve 1.12 kat daha yüksek hücre derişimleri elde edilmiştir. Ayrıca, *P. pastoris* X-33 ve *ADH2*-OE'ye kıyasla fermentasyonun t = 36 st'te *ACS1*-OE için hücre dışına aktarılan asetik asit derişimlerinde 1.49 kat ve 1.78 kat azalma olmuştur. *ADH2*-OE + *ACS1*-OE suşları için t = 9 st'te spesifik etanol tüketim hızı (q_{EtOH}) sırasıyla 0.52 and 0.35 g g^{-1} h^{-1} olarak bulunmuştur; *P. pastoris* X-33 (q_{EtOH} = 0.19 $gg^{-1}h^{-1}$) ile karşılaştırıldığında, *ADH2* aşırı ekspresyonunu etanol tüketim hızını artırmıştır. *ADH2* ve *ACS1*'in aşırı ekspresyonunun merkezi karbon metabolizmasının transkripsiyon seviyeleri üzerindeki etkileri araştırılmıştır. Etanol tüketim yolizi üzerinde yer alan üç enzimi kodlayan genler, *ADH2*, *ALD4* ve *ACS1*, metabolik mühendislikle tasarlanmış üç ekspresyon sistemininde de upregüle olmuştur. *ACS1* aşırı ekspresyonunun r-protein üretimi üzerindeki etkisini araştırmak için, raportör r-protein *mApple* geninin $P_{ADH2\text{-}Cat8\text{-}L2}$'nin kontrol altında ekspres edildiği ve *ACS1*'in $P_{AOX1/Cat8\text{-}L3}$ altında aşırı ekspres edildiği yeni bir ekspresyon sistemi geliştirilmiştir. Geliştirilen bu yeni konak hücre, *ACS1*-OE-C4+mApple-C1 ile %1 (h/h) başlangıç etanol derişiminde üretimde t= 24 st'te *mApple* üretiminin 1.2-kat, t=45 st'te ise 1.32-kat arttığı bulunmuştur.

Anahtar Kelimeler: *Pichia pastoris*, Metabolik mühendisliği, Etanol tüketim yolu, Transkripsiyon analizi

To My Lovely Family and Dear Öztuğ

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my supervisor Prof. Dr. Pınar Çalık, for her never-ending support, great encouragement, precious advice, and help throughout my graduate study in her research laboratory in METU. Her supervision and advice not only improved my academic life but also greatly contributed to my personal maturation. I will always keep in mind what I learned from her and her advice.

I am grateful to my co-supervisor, Prof. Dr. H. Tunçer Özdamar, for his valuable comments and guidance during this MSc thesis.

I would like to thank my friends in the Industrial Biotechnology Laboratory group sincerely: Yiğit Akgün, Özge Kalender, İrem Demir, Beste Avcı, and Oğuz Ulaş Yaman, for their endless support and kind friendship. I especially wish to thank Özge Kalender and İrem Demir for their great support and help in the hardest times of this study.

Middle East Technical University Research Fund Project and Scientific and Technical Research Council of Turkey (TÜBİTAK- 116Z215, TÜBİTAK- 119Z435, BIDEB 2210-C) are greatly acknowledged for financial support.

Last but not least, I would like to deeply thank my beloved family for always being there for me and make me stronger. I would never have achieved this success without them. A special thanks go to Aydın Dayankaç, Sema Dayankaç, Seçil Dayankaç Ünver and Deniz Ünver who have always supported and encouraged me throughout my life. I would also like to express my sincere and dear thanks to my wonderful friend, Hilal Bilgiç, for her sincere support, encouragement, and always being there for me under all circumstances. My special thanks go to Öztuğ Öztürk for brightening even the darkest days with his light, for always making me feel his eternal belief, and most importantly, for never letting me forget what I wanted. This is for you.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xv

LIST OF FIGURES xviii

LIST OF ABBREVIATIONS xxvi

CHAPTERS

1 INTRODUCTION 1

2 LITERATURE REVIEW 5

2.1 Host Microorganism Selection 5

2.1.1 *Pichia pastoris* 11

2.2 Promoters 15

2.2.1 Constitutive Promoters 17

2.2.2 Inducible Promoters 18

2.2.3 Engineered Promoters 19

2.3 Target Proteins 21

2.3.1 *eGFP* Protein 24

2.3.2 *mApple* Protein 25

2.4 Ethanol Utilization Pathway of *Pichia pastoris* 26

2.5 Metabolic Engineering 28

2.6 Transcript-level Expression Analysis 31

2.7 Specific Ethanol Consumption Rates in Air-filtered Shake-flask Bioreactors 33
3 MATERIALS AND METHODS 35
3.1 Kits and Enzymes, Chemicals and DNA Ladders 35
3.2 Strains, Plasmids, Primers, and Maintenance 35
3.3 Genetic Engineering Methods 41
3.3.1 Plasmid Isolation 41
3.3.2 Genomic DNA Isolation 41
3.3.3 Agarose Gel Electrophoresis 42
3.3.4 Extraction of DNA Fragments 43
3.3.5 Gel Elution 43
3.3.6 PCR Purification 44
3.3.7 DNA Sequencing 44
3.3.8 Transformation to *E. coli* 45
3.3.9 Transformation to *P. pastoris* 46
3.4 Construction of Strains and Plasmids 47
3.4.1 Construction of Recombinant Plasmids pADH2-wt::*ADH2* 47
3.4.1.1 Primer Design 47
3.4.1.2 Amplification of the desired DNA 48
3.4.1.3 Purification of the PCR Products and Digestion Reactions 49
3.4.1.4 Ligation reaction of the vector and the insert 52
3.4.2 Construction of Recombinant Plasmids pAOX1-Cat8-L3::*ACS1* 59
3.4.2.1 Amplification of the desired DNA 59
3.4.2.2 Purification of the PCR Products and Digestion Reactions 62

3.4.2.3 Ligation reaction of the vector and the insert 66
3.4.3 Construction of Recombinant Plasmids pmAOX1::*mApple*..... 67
3.4.3.1 Amplification of the desired DNA.....67
3.4.3.2 Purification of the PCR Products and Digestion Reactions.....71
3.4.3.3 Ligation reaction of the vector and the insert 72
3.5 Determination of the gene-copy number.....75
3.6 Growth Media 80
3.6.1 Solid Medium 80
3.6.2 Precultivation Medium 81
3.6.3 Fermentation Media.....82
3.7 Screening Conditions 84
3.7.1 mApple and eGFP synthesis from selected r-*P. pastoris* strains.....84
3.8 Analyses 85
3.8.1 Cell concentration.....85
3.8.2 mApple and eGFP synthesis.....85
3.8.3 Ethanol concentration measurements 86
3.8.4 Extracellular metabolite concentration 86
3.9 Total RNA isolation 87
3.9.1 Transcript level analysis 89
4 RESULTS AND DISCUSSION 93
4.1 Effect of ethanol on EUT pathway.....94
4.2 Overexpression of the enzymes in the EUT pathway 95
4.2.1 Overexpression of Alcohol dehydrogenase 2 (ADH2) enzyme 95
4.2.1.1 Construction of the recombinant plasmid and strains carrying pADH2-wt::*ADH2* 96

4.2.1.2 Influences of *ADH2* overexpression on the growth of *P. pastoris* strains on the carbon sources ethanol and glucose.....112

4.2.2 Overexpression of Acetyl-CoA synthetase 1 (ACS1) enzyme.....116

4.2.2.1 Construction of the recombinant plasmid and strains carrying pAOX1/Cat8-L3::*ACS1* 116

4.2.2.2 Gene-copy number determinations of *ADH2*, *ACS1*, and *mApple* in selected *P. pastoris* Strains 126

4.2.2.3 Influences of *ACS1* overexpression on the growth of *P. pastoris* strains on carbon source ethanol 135

4.3 Transcript level analysis of selected central carbon metabolism genes in the generated recombinant strains 140

4.4 Effects of *ACS1* overexpression on r-protein production on ethanol.....153

4.5 *P. pastoris* strains with double-promoter expression system architectures .157

4.5.1 Recombinant plasmids constructed with DPES and SPES in *P. pastoris* 158

4.5.2 Screening of Constructed DPESs and SPESs.....165

5 CONCLUSIONS.....171

REFERENCES.....175

APPENDICES

A. Promoter and gene sequences 187

B. Nucleotide sequences of constructed plasmids.....200

C. Buffers and Stock Solutions.....205

D. DNA Markers.....207

E. Standard Curves for Organic Acids for HPLC Analysis 208

LIST OF TABLES

TABLES

Table 2. 1 *Comparison of common organisms used as a host in industrial production (Gustavsson, 2018), (Vogl* et al., *2013), (Berlec and Strukelj, 2013), (Demain* et al., *2009), (Mattanovich* et al., *2012)* 8

Table 2. 2 *Properties of most commonly used species in bioprocesses (Gustavsson, 2018) (Mattanovich* et al., *2012)* 10

Table 2. 3 *Advantages and disadvantages of P. pastoris expression system* 13

Table 2. 4 *Constitutive promoters of P. pastoris (Çalık* et al., *2015) (Massahi and Çalık, 2018)* 18

Table 2. 5 *Constitutive promoters of P. pastoris (Çalık* et al., *2015)* 19

Table 2. 6 *Some common metabolic engineering strategies* 31

Table 3. 1 *Constructed strains and plasmids*…….. 36

Table 3. 2 *Designed primers that were used in this study.* 39

Table 3. 3 *PCR operation conditions for amplification of P_{ADH2}-ADH2*.....48

Table 3. 4 *Composition of PCR reaction mixture for P_{ADH2}-ADH2* 49

Table 3. 5 *Conditions of single digestion reaction with KpnI for pGAPZαA* 51

Table 3. 6 *Conditions of single digestion reaction with BglII for pGAPZαA*.....51

Table 3. 7 *Conditions of double digestion reaction with BamHI and KpnI for the insert* 52

Table 3. 8 *Composition of the ligation reaction mixture of pGAPZαA and ADH2-ADH2*.....53

Table 3. 9 *PCR operation conditions for amplification of hygromycin* 55

Table 3. 10 *Composition of PCR reaction mixture for hygromycin*.....55

Table 3. 11 *Conditions of double digestion reaction with BamHI and PciI for vector plasmid* 56

Table 3. 12 *Conditions of double digestion reaction with BamHI and PciI for the insert* 57

Table 3. 13 *Composition of the ligation reaction mixture of pADH2:: ADH2 and Hygromycin* 58

Table 3. 14 *PCR operation conditions for amplification of pAOX1-Cat8-L3*.....59
Table 3. 15 *PCR operation conditions for amplification of ACS1* 60
Table 3. 16 *Composition of PCR reaction mixtures for pAOX1-Cat8-L3 and ACS1*....60
Table 3. 17 *Thermo-cyclic PCR operation conditions for construction of* $P_{AOX1\text{-}Cat8\text{-}L3}$-*ACS1*.....62
Table 3. 18 *Composition of PCR reaction mixture for construction of* $P_{AOX1\text{-}Cat8\text{-}L3}$-*ACS1*.....62
Table 3. 19 *Conditions of single digestion reaction with KpnI for pGAPZαA*.....64
Table 3. 20 *Conditions of single digestion reaction with BglII for pGAPZαA*.....65
Table 3. 21 *Conditions of double digestion reaction with BamHI and KpnI for the insert*.....65
Table 3. 22 *Composition of the ligation reaction mixture of vector and insert* 66
Table 3. 23 *PCR operation conditions for amplification of pmAOX1*.....68
Table 3. 24 *PCR operation conditions for amplification of mApple*.....68
Table 3. 25 *Composition of PCR reaction mixtures for pmAOX1 and mApple* 69
Table 3. 26 *Thermo-cyclic PCR operation conditions for construction of pmAOX1-mApple*.....70
Table 3. 27 *Composition of PCR reaction mixture for construction of* P_{mAOX}-*mApple*.70
Table 3. 28 *Composition of the digestion reaction mixture with AscI and XbaI for vector, pADH2-Cat8-L2::hGH*.....71
Table 3. 29 *Composition of the digestion reaction mixture with AscI and XbaI for insert, PmAOX1::mApple*.....72
Table 3. 30 *Composition of the ligation reaction mixture of vector and insert* 73
Table 3. 31 *Primers designed for the qPCR experiments and amplification of standards* 75
Table 3. 32 *Composition of PCR reaction mixture for standards*.....78
Table 3. 33 *PCR operation conditions for amplification of standards* 78
Table 3. 34 *Composition of qPCR reaction mixture* 79
Table 3. 35 *Thermocycler operation-profile of qPCR experiments* 80
Table 3. 36 *Composition of LB Agar Medium for E. coli*.....81
Table 3. 37 *Composition of YPD Agar Medium for P. pastoris*.....81

Table 3. 38 *Composition of the BMGY for the precultivation of P. pastoris cells* 82
Table 3. 39 *Composition of the YP for the precultivation of P. pastoris cells* 82
Table 3. 40 *Defined fermentation medium for screening* 83
Table 3. 41 *ASMV6 Defined cultivation-base-fermentation medium for screening*.....83
Table 3. 42 *List of primers used for the transcript level analysis.* 89
Table 4. 1 *PCR operation conditions for colony PCR*…… 101
Table 4. 2 *Composition of PCR reaction mixture for colony PCR*.....102
Table 4. 3 *Nanodrop measurement results of genomic DNA samples isolated from different P. pastoris strains* 128
Table 4. 4 *Nanodrop results of standards* 130
Table 4. 5 *Standard curve results of ARG4*.....131
Table 4. 6 *Standard curve results of ADH2* 132
Table 4. 7 *Standard curve results of ACS1*.....132
Table 4. 8 *Standard curve results of mApple* 132
Table 4. 9 *Standard curve data of ARG4, ADH2, ACS1, and mApple* 132
Table 4. 10 *Relative quantification results of selected P. pastoris strains.*.....133
Table 4. 11 *Constructed P. pastoris strains, gene copy numbers, and their abbreviations*.....134
Table 4. 12 *Quality and Quantity of P. pastoris RNA Samples Prepared by RNaesy Mini Kit (Qiagen)* 140
Table 4. 13 *Genes involved in central carbon metabolism of P. pastoris* 142

LIST OF FIGURES

FIGURES

Figure 2. 1 Percentage of biopharmaceuticals produced in different expression systems (Baeshen *et al.,* 2014) 7
Figure 2. 2 Central carbon metabolism of *Pichia pastoris* (Çalık et al., 2015) 14
Figure 2. 3 Structure of a promoter and eukaryotic gene transcription 16
Figure 2. 4 Main areas of applications of fluorescent proteins (Chudakov et al., 2010) 22
Figure 2. 5 Spectral diversity of available monomeric FPs. 23
Figure 2. 6 A diagram of ethanol utilization pathway and TCA cycle in *Pichia pastoris* 28
Figure 2. 7 Strategy of metabolic engineering for the production of a desired chemical: overexpression of therate-limiting enzyme. 29
Figure 2. 8 Strategy of metabolic engineering for the production of a desired chemical: inhibition of the competing pathway. 29
Figure 3. 1 Map of the constructed plasmid (pADH2:: *ADH2*) from pGAPZα 54
Figure 3. 2 Map of the constructed plasmid (pADH2:: *ADH2*) with a hygromycin resistance 58
Figure 3. 3 Algorithm of construction logic of $P_{AOX1\text{-}Cat8\text{-}L3}$::*ACS1* fragment 61
Figure 3. 4 Map of the constructed plasmid pAOX1-Cat8-L3:: *ACS1* from pGAPzα-A. 67
Figure 3. 5 Algorithm of construction logic of P_{mAOX1}::mApple fragment 69
Figure 3. 6 Map of the constructed plasmid pmAOX1:: mApple from pADH2-Cat8-L2::*hGH* 73
Figure 3. 7 Schematic representation of the metabolic engineering strategy for construction of desired plasmid 74
Figure 4. 1 By-product concentrations in the cell culture at t = 36 h excreted from wild-type *P. pastoris* X-33 strain on 1% (v/v) and 2% (v/v) ethanol 94
Figure 4. 2 Ethanol utilization pathway in *Pichia pastoris* 95

Figure 4. 3 Base plasmid pGAPZα-A, and restriction enzymes *Bgl*II and *Kpn*I, used in construction of recombinant plasmid. 96
Figure 4. 4 Schematic representation of restriction endonucleases used in the construction of pADH2-wt::*ADH2* plasmid 97
Figure 4. 5 Agarose gel electrophoresis image of the gene sequences amplified with Forward_$P_{ADH2\text{-}wt}$ and Reverse_*ADH2* primers at different annealing temperatures. 1: Gene Ruler Express (Thermoscientific); Annealing temperatures: 2: T= 60.8°C; 3: T = 62.7° C; 4: T = 65.7° C; 5: T = 67.8° C; 6: T = 69.2° C 98
Figure 4. 6 Agarose gel electrophoresis image of the purified DNA fragment of pADH2-wt::*ADH2*. 1: Quick-Load Purple 1 kb DNA ladder; 2: Purified DNA fragment of pADH2-wt::*ADH2* 99
Figure 4. 7 Agarose gel electrophoresis of double digested pADH2-wt::*ADH2* fragment and plasmid pGAPZαA 1: Quick-Load Purple 1 kb DNA ladder; 2: Digested insert with *Bam*HI and *Kpn*I; 3: Digested base-plasmid, pGAPZαA, with *Kpn*I and *Bgl*II. 100
Figure 4. 8 Agarose gel electrophoresis of purified double digested pADH2-wt::*ADH2* fragment and plasmid pGAPZαA 1: GeneRuler Express; 2: Digested insert with *Bam*HI and *Kpn*I after gel elution; 3: Digested base-plasmid, pGAPZαA, with *Kpn*I and *Bgl*II REs after gel elution. 100
Figure 4. 9 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. M1: GeneRuler Express; 1: Positive control for PCR control of insert pADH2-*ADH2* gene; 2: Negative control for PCR control of insert pADH2-*ADH2* gene, 3-8: PCR products of potential plasmids.....102
Figure 4. 10 Agarose gel electrophoresis of potential recombinant plasmids. 1: Generuler Express, 2-7: Potential plasmids carrying the $P_{ADH2\text{-}wt}$::*ADH2* fragment ...103
Figure 4. 11 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. M1: GeneRuler Express; 1: Positive control for PCR control of insert pADH2-*ADH2* gene; 2: Negative control for PCR control of insert pADH2-*ADH2* gene, 3-8: PCR products of potential plasmids.....103

Figure 4. 12 Agarose gel electrophoresis of purified double digested Hygromycin with TEF promoter and TEF terminator fragment 1: GeneRuler Express; 2: Digested insert with *Bam*HI and *Pci*I REs after PCR purification. 105
Figure 4. 13 Agarose gel electrophoresis image of 1: Double digested vector with *Bam*HI and *Pci*I; 2: GeneRuler Express. 105
Figure 4. 14 Agarose gel electrophoresis of double digested vector-plasmid pADH2-wt::*ADH2* 1: GeneRuler Express; 2:Digested and purified vector-plasmid with *Bam*HI and *Pci*I.....106
Figure 4. 15 Agarose gel electrophoresis of potential *E. coli* clones carrying the Hygromycin gene after transformation. 1: GeneRuler Express; 2: Negative control for PCR control of insert Hygromycin gene; 2: Negative control for PCR control of insert Hygromycin gene, 4-15: PCR products of potential colonies. 107
Figure 4. 16 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. 1: GeneRuler Express; 2: Negative *control for PCR control of insert* pADH2-wt*::ADH2 gene; 2: Negative control for PCR control of insert* pADH2-wt*::ADH2 gene, 4-15: PCR products of potential colonies.*.....108
Figure 4. 17 Agarose gel electrophoresis of potential recombinant plasmids. 1: Generuler Express, 2-5: Potential plasmids carrying the Hygromycin and pADH2-wt::*ADH2* cassettes.....108
Figure 4. 18 Agarose gel electrophoresis of control PCR for potential recombinant plasmids. 1:Generuler Express; 2-5: Control PCR with Forward_Hyg and Reverse_Hyg primers. (Hygromycin gene 1479bp).....109
Figure 4. 19 Agarose gel electrophoresis of control PCR for potential recombinant plasmids. 1:Generuler Express; 2-5: Control PCR with pADH2_forward and ADH2_reverse primers. (pADH2-wt::*ADH2* cassette 2100 bp) 110
Figure 4. 20 Agarose gel electrophoresis of linearized plasmid pADH2-wt::*ADH2* plasmid carrying hygromycin fragments 1: pADH2-wt::*ADH2* plasmid after linearization with *Bam*HI ; 2: Generuler Express.....111
Figure 4. 21 Schematic representation of constructed *P. pastoris* expression system. 111

Figure 4. 22 Variations in the cell concentrations (DCW) of the strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* overexpression cassettes in batch fermentation of 2 % (v/v) ethanol, and excess glucose (20 g/L) and for comparison with *P. pastoris* X-33, with the cultivation time in shake-bioreactors in V_R = 50-mL fermentation volume at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}.....113
Figure 4. 23 By-product concentrations in the cell culture at t = 36 h for *ADH2* overexpression strains and wild-type *P. pastoris* X-33 strain on 2% (v/v) ethanol 114
Figure 4. 24 Base plasmid pGAPZα-A (KanMX), and the restriction enzymes *Bgl*II and *Kpn*I, used in construction of recombinant plasmid. 117
Figure 4. 25 Agarose gel electrophoresis of the gene sequences amplified with Forward_ACS1and Reverse_ACS1 primers at different annealing temperatures. 1: Quick-Load Purple 1 kb DNA ladder; Annealing temperatures: 2:T= 64.3°C; 3: T= 66.5° C.....118
Figure 4. 26 Agarose gel electrophoresis image of the gene sequences amplified with Forward_pAOX1-Cat8-L3and Reverse_pAOX1-Cat8-L3 primers at different annealing temperatures. 1: G Quick-Load Purple 1 kb DNA ladder; Annealing temperatures: 2: T= 64.3° C; 3: T= 66.5° C. 119
Figure 4. 27 Agarose gel electrophoresis image of the purified DNA fragments of pAOX1-Cat8-L3 and *ACS1*. 1&3: Quick-Load Purple 1 kb DNA ladder; 2: Purified DNA fragment of *ACS1* 4: Purified DNA fragment of pAOX1-Cat8-L3. 120
Figure 4. 28 Agarose gel electrophoresis of double digested vector, pGAPZA , and insert, pAOX-Cat8-L3::*ACS1*. 1: Digested insert with *Bam*HI and *Kpn*I; 2: Quick-Load Purple 1 kb DNA ladder; 3: Digested base-plasmid, pGAPZαA-KanMX, with *Kpn*I and *Bgl*II. (Before gel extraction).....121
Figure 4. 29 Agarose gel electrophoresis of purified double digested vector, pGAPZαA, and insert, pAOX-Cat8-L3::*ACS1*.1: Quick-Load Purple 1 kb DNA ladder; 2: Digested insert with *Bam*HI and *Kpn*I REs after gel elution; 3: Digested base-plasmid, pGAPZαA-KanmMX, with *Kpn*I and *Bgl*II REs after gel elution. 122
Figure 4. 30 Agarose gel electrophoresis image of potential *E. coli* clones carrying the $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* gene after transformation. M1: Quick-Load Purple 1 kb DNA ladder; 1: Positive control for PCR control of insert $P_{AOX\text{-}Cat8\text{-}L3}$-*ACS1* gene; 2: Negative

control for PCR control of insert $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* gene, 3-14: PCR products of potential plasmids.....123
*Figure 4. 31*Agarose gel electrophoresis image of potential recombinant plasmids. 1: Quick-Load Purple 1 kb DNA ladder, 2-5: Potential plasmids carrying the $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* Another colony PCR was performed with Forward_pAOX1-Cat8-L3 and Reverse_ACS1 primers under the conditions described in Tables 4.1 and 4.2, using isolated plasmids as a template. (Figure 4.32) 123
Figure 4. 32 Agarose gel electrophoresis of potential *E. coli* clones carrying the pAOX1-Cat8-L3::*ACS1* gene after transformation. M1: Quick-Load Purple 1 kb DNA ladder; 1: Positive control for PCR control of insert pAOX-Cat8-L3-*ACS1* gene; 2: Negative control for PCR control of insert pAOX-Cat8-L3-*ACS1* gene, 3-7: PCR products of potential plasmids.....124
Figure 4. 33 Agarose gel electrophoresis image of 1&3: pAOX1-Cat8-L3::*ACS1* plasmid after linearization with *Bam*HI RE; 2: Quick-Load Purple 1 kb DNA ladder. 125
Figure 4. 34 Schematic representation of the constructed *P. pastoris* expression systems. 126
Figure 4. 35 Agarose gel electrophoresis of isolated genomic DNA of selected *P. pastoris* strains 1: Quick-Load Purple 1 kb DNA ladder; 2-7: Isolated Genomic DNA for selected colonies (600-1, 600-2, 600-5, 600-6, 750-1, 750-2) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1*+ $P_{ADH2\text{-}Cat8\text{-}L2}$::*mApple* expression system. 127
Figure 4. 36 Agarose gel electrophoresis of isolated genomic DNA of selected *P. pastoris* strains 1: Quick-Load Purple 1 kb DNA ladder; 2-3: Isolated Genomic DNA for selected colonies (12-13) of *P. pastoris* strains that contain $P_{ADH2\text{-}wt}$::*ADH2* plasmid; 5: Isolated Genomic DNA for the selected colony (500-1) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1*; 6: Isolated Genomic DNA for the selected colony (600-2) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1*+ $P_{ADH2\text{-}wt}$::*ADH2* expression system. 127
Figure 4. 37 Agarose gel electrophoresis of *ARG4*, *ADH2* and *ACS1* genes amplified by PCR 1: Quick-Load Purple 1 kb DNA ladder; 2: Amplified standard *ARG4* gene; 3: Amplified standard *ADH2* gene; 4: Amplified standard *ACS1* gene.....129

Figure 4. 38 Agarose gel electrophoresis of mApple gene amplified by PCR 1: Quick-Load Purple 1 kb DNA ladder; 2-5: Amplified *mApple* gene as standard. 130
Figure 4. 39 Variations in the cell concentrations (DCW) of strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$::ACS1 overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33, with the cultivation time in shake-bioreactors in V_R = 50 mL fermentation volume at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}.....135
Figure 4. 40 Variations in ethanol concentrations of strains constructed with $P_{ADH2\text{-}wt}$*::ADH2* and $P_{AOX1/Cat8\text{-}L3}$:: *ACS1* overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33 with cultivation time 137
Figure 4. 41 Variations in the specific ethanol uptake rates of strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$:: *ACS1* overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33 with cultivation time 137
Figure 4. 42 By-product concentrations in the cell culture at t=36 h for overexpression strains and wild-type *P. pastoris* X-33 strain on 1% (v/v) ethanol 139
Figure 4. 43 Central carbon metabolism pathways of *P. pastoris* (Paes *et al.,* 2021) 143
Figure 4. 44 Influence of *ADH2* overexpression on transcript levels of selected genes in *P. pastoris* induced on ethanol 144
Figure 4. 45 Influence of *ADH2* and *ACS1* overexpression on transcript levels of selected genes in *P. pastoris* induced on ethanol 147
Figure 4. 46 The gene expression ratios (ER) based on housekeeping gene *ACT1* obtained by *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE + *ACS1*-OE in central metabolism. 148
Figure 4. 47 Venn diagrams of upregulation ratios of genes in *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE+*ACS1*-OE strains. (A) Upregulation of genes involved in C2 metabolism; (B) Upregulation of genes involved in TCA cycle; (C) Upregulation of genes involved in gluconeogenesis. 150
Figure 4. 48 Organic acid concentrations in the cell culture at t = 36 h for wild-type *P. pastoris X-33*, *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE+*ACS1*-OE overexpression strains on 1% (v/v) ethanol 152

Figure 4. 49 mApple expressions of *P. pastoris* strains constructed with $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* and *ACS1* overexpression cassettes at t = 24 h of batch fermentation in ASMv6 medium on the carbon source of ethanol in V_R = 2-mL cultures in 12-deep well plates at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}. Reporter FP mApple expressed under the EPV *ADH2-Cat8-L2*. Expressions with recombinant host cells and overexpression strains at 0.5 % (*v/v*) and 1 % (*v/v*), of ethanol. Error bars represent the standard deviation (±). 154
Figure 4. 50 Variations in the *mApple* expression levels with the cultivation time in air-filtered shake bioreactors on 1% (v/v) ethanol. 156
Figure 4. 51 Variations in specific ethanol uptake rates with the cultivation time in air-filtered shake bioreactors on 1% (v/v) ethanol. 156
Figure 4. 52 Base plasmid, pADH2-Cat8-L2::*hGH*, and restriction enzymes, *Asc*I and *Xba*I, used in construction of recombinant plasmid. 159
Figure 4. 53 Agarose gel electrophoresis of the gene sequences amplified with different primer combinations to construct pmAOX::*mApple* plasmid. 1: Quick-Load Purple 1 kb DNA ladder 2: P_{mAOX}_Forward and P_{mAOX}_Reverse; 3: *mApple*_Forward and mApple_Reverse.....160
Figure 4. 54 Agarose gel electrophoresis of the gene sequences amplified with Forward_P_{mAOX} and Reverse_mApple to construct pmAOX::*mApple* fragment. 1: Quick-Load Purple 1 kb DNA ladder; 2-5: Amplified pmAOX::*mApple* fragment. ..161
Figure 4. 55 Agarose gel electrophoresis of 1: Double digested vector with *Asc*I and *Xba*I; 2: GeneRuler Express; 3: Double digested insert, pmAOX::*mApple*, with *Asc*I and *Xba*I 161
Figure 4. 56 Agarose gel electrophoresis of potential *E. coli* clones carrying the P_{mAOX}::*mApple* gene after transformation. M1: Quick-Load Purple 1 kb DNA ladder; 1: Negative control for PCR control of insert pADH2-*ADH2* gene; 2: Positive control for PCR control of insert pADH2-*ADH2* gene, 3-6: PCR products of potential plasmids 162
Figure 4. 57 Agarose gel electrophoresis image of potential recombinant plasmids. 1: Quick-Load Purple 1 kb DNA ladder, 2-5: Potential plasmids carrying the $P_{ADH2\text{-}wt}$::*ADH2* gene.....163

Figure 4. 58.....164
Figure 4. 59 mApple expression levels of *P. pastoris* strains carrying pmAOX::*mApple* on 1% (v/v) ethanol at t= 24 h of the fermentation. Error bars represent the standard deviation. 166
Figure 4. 60 *mApple* and *eGFP* expression levels of *P. pastoris* strains carrying the DPES pmAOX::*mApple* + pGAP::*eGFP* on 1% (v/v) ethanol at t=24 h of the fermentation. *mApple* and *eGFP* represent the expression levels of the constituent NEPV ADH2-Cat8-L2 and the NOP GAP, respectively. 167
Figure 4. 61 Variations in the *mApple* and *eGFP* expression levels with the cultivation time in the air-filtered shake bioreactors on 20 g/L glucose. 168
Figure E1 Standard curve for acetaldehyde…….. 208
Figure E2 Standard curve for acetic acid 208
Figure E3 Standard curve for oxalic acid.....209
Figure E4 Standard curve for malic acid 209
Figure E5 Standard curve for glyoxylic acid 210
Figure E6 Standard curve for pyruvic aci 210

LIST OF ABBREVIATIONS

ACS1	Acetyl-CoA synthetase	
ADH2	Alcohol dehydrogenase 2	
AOX1	Alcohol oxidase 1	
ARG4	Argininosuccinate lyase	
BLAST	Basic local alignment search tool	
bp	Base pair	
C	Concentration in the medium	$g\ L^{-1}$
EDTA	Ethylene diamine tetra acetic acid	
GAP	Glyceraldehyde-3-phophate dehydrogenase	
GC	Gas chromatography	
GRAS	Generally regarded as safe	
HPLC	High pressure liquid chromatography	
LB	Luria Broth	
NEPV	Novel engineered promoter variant	
OD_{600}	Optical density at 600 nm	
OE	Overexpression	
P	Promoter	
PCR	Polymerase chain reaction	
PTM1	Pichia trace minerals medium	
q	Specific consumption rate	

rDNA	Recombinant DNA	
t	Cultivation time	h
T	Bioreaction medium temperature	
TCA	Tricarboxylic acid	
TE	Tris-EDTA	
TF	Transcriptional factor	
V	Volume of the bioreactor	

CHAPTER 1

INTRODUCTION

Metabolic engineering is the directed modulation of metabolic pathways for metabolite over production or the improvement of cellular properties. It uses different databases, libraries of components and conditions to generate the maximum production rate of a desired chemical compound and avoiding inhibitors and conditions that affect the growth rate and other vital functions in the specific organism to achieve these goals; metabolic fluxes manipulation represents an important alternative (Stephanopoulos, 2012). Systems metabolic engineering integrates tools and strategies of systems biology, synthetic biology, and evolutionary engineering with traditional metabolic engineering, and it has recently been used to facilitate the development of high-performance strains (Choi et al., 2019). Promoter engineering is an important metabolic engineering tool to generate engineered-promoters that are the key components in the cell-factory design, allowing precise and enhanced expression of genes. Promoters having exceptional strength are attractive candidates for designing metabolic engineering strategies by changing the induction mechanism of promoters to take advantage of alternative metabolic pathways.

Yeasts have become prominent in heterologous expression due to their efficient expression capacity of eukaryotic proteins and the ability of proper folding/modification of the product for full activity (Çelik and Çalık, 2012). Yeast metabolic engineering has largely focused on assembling metabolic pathways in the cytosol. Strategies to spatially optimize metabolic pathways, such as protein fusions and synthetic protein scaffolds, have effectively boosted the productivity of some engineered pathways (Dueber et al., 2009).

Among yeast cells, *Pichia pastoris* is qualified as an ideal host for the production of many r-proteins with its ability to produce r-proteins both intracellularly and extracellularly without the need of many purification steps, and most importantly, it has tightly regulated expression systems that become the reasons for preference. Metabolic engineering in *P. pastoris* has been conducted on the cellular machineries, central carbon metabolism, redox and energy generation (Kafri et al., 2016, Klein et al., 2015, Wu et al., 2016, Pena et al., 2018). The production of the biofuel isobutanol from two renewable carbon sources, glucose and glycerol, was increased by exploiting the yeast's amino acid biosynthetic pathway and diverting the amino acid intermediates to the 2-keto acid degradation pathway (Siripong et al., 2018). By increasing the ADC copy number and C4 precursor supply via the overexpression of aspartate dehydrogenase, feasible β-alanine production from methanol with optimization of aspartate decarboxylation was also achieved (Miao et al., 2021). Furthermore, by the overexpression of *GND2, SOL3* , and *RPE1* the flux towards the pentose phosphate pathway (PPP), human interferon-gamma (hIFN-γ) production in *P. pastoris* was increased (Prabhu and Veeranki, 2019) and by integrating the D-xylulose-forming D-arabitol dehydrogenase (*DalD*) gene from *Klebsiella pneumoniae* and the xylitol dehydrogenase (*XDH*) gene from *Gluconobacter oxydans* into *P. pastoris* GS225, recombinant strains were constructed to convert xylitol from glucose by the pathway of glucose–D-arabitol–D-xylulose–xylitol in a single process; as a result, conversion of glucose to xylitol with the highest yield was achieved (Cheng et al., 2014). An engineered dynamically controlled yeast biosynthetic system that allows continuous polyketide synthesis but automatically controlled competitive and essential FAs synthesis was generated to improve biosynthesis of malonyl-CoA derived bioactive chemicals. The optimal distribution of malonyl-CoA flux to the target compounds was proved by the improved polyketide production, and this new system provides an applicable method for efficient synthesis of malonyl-CoA derived compounds in eukaryotic chassis hosts (Wen et al., 2020). On the other hand, several metabolic engineering approaches have been applied to increase the production capacity of *S. cerevisiae*;

for instance, to increase cellulase production by decreasing the metabolic burdens by POT1-mediated δ-integration, metabolic engineering strategies, including optimization of codon usage, promoter, and the signal peptide was applied (Song et al., 2018). Furthermore, increased succinic acid production was achieved with the elimination of the ethanol biosynthesis pathways by introducing the *mae1* gene encoding the *Schizosaccharomyces pombe* malic acid transporter into *S. cerevisiae* (Ito et al., 2014).

As the selection of efficient and tunable promoters is very critical for r-protein production, several constitutive and inducible *P. pastoris* promoters have been identified with distinct properties and strengths (Massahi and Çalık, 2018). One of the most commonly used inducible promoters of *P. pastoris* is the strong *Alcohol oxidase* 1 (P_{AOX1}) promoter, the gene product of which catalyzes the first reaction in methanol utilization pathway, conversion of methanol to formaldehyde, of *P. pastoris*. One other inducible promotor of *P. pastoris* is the *Alcohol dehydrogenase 2* (P_{ADH2}) promoter, the gene product of which catalyzes the reaction that converts ethanol to acetaldehyde (Cregg & Tolstorukov, 2012).

For pharmaceuticals and food industries, since the use of methanol as the substrate in r-protein productions may complicate: ***i)*** separation processes, and ***ii)*** required documentations, methanol-free production by green-clean-bioprocesses is highly encouraged and challenging through discovery and engineered-novel promoter variants (NEPVs) and systems for novel r-*P. pastoris* strains construction. As the yield and productivity in methanol-free production is an industrial target, to compete with methanol-based production under P_{AOX1}, the discovery of strong promoters-, especially, design and construction of strong NEPVs and systems to be active in alternative pathways, is required. For this purpose, strong ethanol-induced *AOX1* and *ADH2* promoter variants were designed and constructed (Ergün *et al.*, 2019, 2020). With the discoveries of strongly ethanol-inducible NEPVs, the utilization of ethanol as a carbon source gains importance. For this reason, the application of metabolic engineering methods to increase the level of gene expressions responsible for ethanol utilization is a promising strategy in bioprocess

applications to increase the production capacity of the host and efficiency. An engineered *Escherichia coli* strain which could utilize ethanol as the sole carbon source was constructed by introducing the alcohol dehydrogenase and aldehyde dehydrogenase from *Aspergillus nidulans* into *E. coli* and the recombinant strain acquired the ability to grow on ethanol (Liang et al., 2021). In *Pichia pastoris*, using metabolic engineering methods, an alternative method for the generation of acetyl-CoA preferentially in cytosol based on the strong ethanol assimilative capacity was proposed. With the synthetic ethanol-induced transcriptional signal amplification device (ESAD) system, ethanol was made serving as a strong expression inducer in addition to its roles of the sole carbon source and acetyl-CoA precursor, and highly improved the production of the acetyl-CoA (Liu et al., 2019).

Alcohol dehydrogenase 2 (ADH2) enzyme is the first enzyme in the EUT pathway that catalyzes the oxidation of ethanol to acetaldehyde. Furthermore, it is experimentally revealed that *ADH2* is the only gene responsible for the consumption of ethanol in *P. pastoris*. (Karaoglan et al., 2016; Ergün et al., 2019). As a result, with metabolic engineering approaches, increasing EUT pathway fluxes is the right strategy to increase r-protein production in P_{ADH2} control.

The ultimate aim of this thesis is designing a novel metabolically engineered host, *P. pastoris*, by overexpression of the gene(s) of the enzyme(s) that catalyzes the reactions in the EUT pathway under a hybrid-architectured synthetic promoter to increase the ethanol uptake rate and decrease by-product formation in the ethanol utilization (EUT) pathway and acquire a smoothly operating intracellular reaction network from the TCA cycle to the gluconeogenesis pathway through anaplerotic reactions.

CHAPTER 2

LITERATURE REVIEW

2.1 Host Microorganism Selection

The discovery of recombinant DNA technology led to the development and manufacturing of a large number of proteins with the therapeutic activity known as biology-based therapeutics. Biosynthetic insulins, peptide hormones, interferons, monoclonal antibodies are examples of protein biopharmaceuticals (Conner *et al*., 2014). Furthermore, these products are very efficient in treating severe acute and chronic diseases such as hormone-related disorders, cancer, and immunologic diseases (Tsiftsoglou *et al.,* 2013). Thus, in recent decades, biotechnology has become one of the most valuable components of both the pharmaceutical and chemical process industries (Fan *et al.,* 2015). Biopharmaceuticals are an integral part of modern medicine and pharmacy. Both the development and the biotechnological production of biopharmaceuticals are highly cost-intensive and require suitable expression systems (Vogl-Glieder *et al.,* 2013). To be able to achieve high yield and productivity in recombinant processes, some critical requirements should be satisfied. In this context, one of the most crucial steps in bioprocess design is the host organism selection since the success or failure of a process often hinges on the initial choice of a host organism and expression system (Liu et al., 2020). Rapid growth on cheap media, moderate process conditions, suitability for scaling-up, and minimum production of the by-products are the critical properties of an ideal host. Additionally, the production of the toxic compounds during the process is an undesired situation, so the ideal host should also satisfy this condition and should be GRAS (Generally Recognized As Safe) status. Another crucial parameter for host selection is the capability of performing post-translational modifications, which are

changes to the proteins after their amino acid sequence have been performed to fold the protein in its correct active form if the target protein requires (Gustavsson, 2018, p. 8).

A large number of host systems, which include bacteria, yeast, insect cells, transgenic animals as well as mammalian and human cell lines, developed over the years with the help of recombinant protein technology, enabled the expression of recombinant protein-based pharmaceuticals (Tsiftsoglo et al., 2013; Tripathi and Shrivastava, 2019). Among these expression systems, *E. coli* and Mammalian cells were more dominant than other microorganisms for the production of recombinant proteins (Figure 2.1) (Ferrer-Miralles et al., 2009). However, mammalian cell culture processes are relatively slow, require complex media, and are susceptible to viral contaminations (Vogl et al., 2013). For this reason, microorganisms can be designated as ideal hosts for r-protein production compared to mammalian or insect cells since they have vital properties for achieving optimum and simple bioprocess conditions, such as enabling easy purification of the product and high growth rates. *Escherichia coli* is one of the most preferred expression systems for r-protein production among host organisms since it is a genetically well-characterized organism, and also its ability of fast growth in inexpensive media and reaching high densities in a short time make this species highly desirable for industrial protein production. (Carter, 2011; Corchero et al., 2013). In spite of its advantages, there are several drawbacks of *E. coli*, which limit its application in r-protein production. One of the most significant drawbacks is its inability to perform post-translational modifications, which are generally necessary for correct protein folding and activity. Post-translational modifications, PTMs, contain glycosylation, disulfide bond formation, phosphorylation, and proteolytic processing, and among these, the most important one is glycosylation since approximately 70% of the therapeutic proteins currently in clinical trials are glycosylated (Sethuraman and Stadheim, 2006; Walsh and Jefferis, 2016). Another significant disadvantage of *E. coli* is the lack of a general secretory system, which causes inefficient protein secretion. Furthermore, other obstacles, such as lack of proper protein folding and inclusion body (IB)

formation, which results in high purification steps, also limit the usage of *E. coli* as a host organism in bioprocesses. (Tripathi and Shrivastava, 2019; Ferrer-Miralles et al., 2009; Corchero et al., 2013).

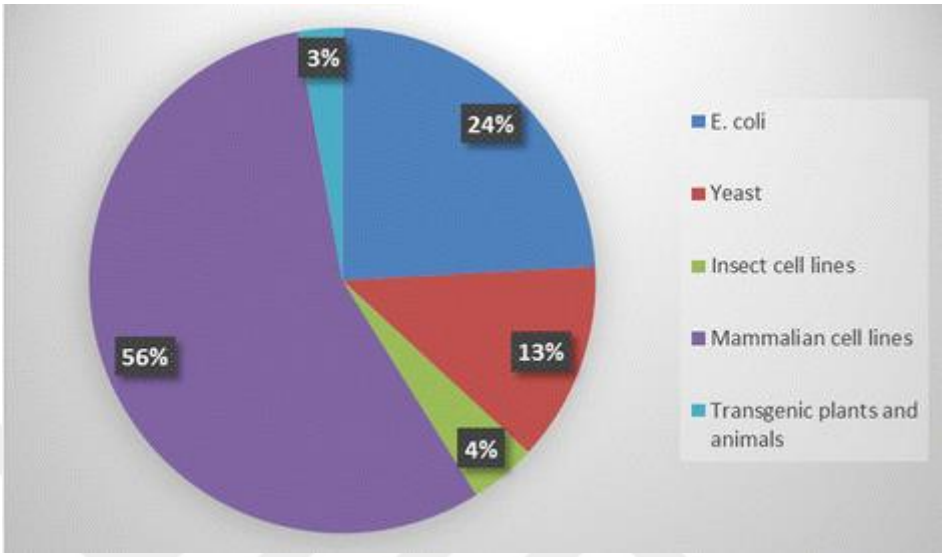


Figure 2. 1 Percentage of biopharmaceuticals produced in different expression systems (Baeshen *et al.,* 2014)

As well as *E. coli*, another most preferred microorganism for the r-protein production is yeast. In particular, when the target protein is a non-soluble form in the prokaryotic system or requires a specific PTM, yeasts become prominent as the host organism (Jenkins, 2007). Furthermore, some significant advantages, such as providing cost-effective bioprocesses and high densities, enabling genetic manipulations easily, and performing some post-translational manipulations and high production capacity of heterologous proteins, make this organism a highly productive cell factory (Ferrer-Miralles et al., 2009). On the other hand, yeasts are not able to perform some complex post-translational modifications such as amidation, prolyl hydroxylation, and some phosphorylation (Çelik and Çalık, 2012).

Table 2. 1 *Comparison of common organisms used as a host in industrial production (Gustavsson, 2018), (Vogl* et al., *2013), (Berlec and Strukelj, 2013), (Demain* et al., *2009), (Mattanovich* et al., *2012)*

	Ease of genetic manipulation	**PTMs**	**Yields**	**Contamination**
Bacteria	Simple	Required additional steps	High expression capacities, inefficient secretion	Possible phage infection
Yeast	Simple	Required additional steps	High yields	Little risk of viral DNAs
Filamentous fungi	Simple	Yes	High yields	Little risk of viral DNAs
Higher Eukaryotes	Moderate	Yes	High yields	Risk of viral contamination

A wide range of outstanding yeast expression hosts, including the species *Saccharomyces cerevisiae, Pichia pastoris*, *Hansenula polymorpha, Kluyveromyces lactis, Schizosaccharomyces pombe, Yarrowia lipolytica, and Arxula adeninivorans*, have been used in recombinant protein production processes with various characteristics such as being thermo-tolerant or halo-tolerant, rapidly reaching high cell densities or utilizing unusual carbon sources (Çelik and Çalık, 2012).

Yeasts can be divided into two categories as crabtree positive and crabtree negative. The Crabtree effect includes energy management in which glycolysis is used by the yeasts as the terminal electron acceptor instead of oxygen, despite sufficiently high dissolved oxygen concentration content, when such concentrations exceed a specific limit (Imura et al., 2020). In other words, some yeast species use fermentation even in the presence of oxygen when glucose concentrations are sufficiently high.

The glucose supply in the culture should be constrained to prevent the crabtree effect, and as a result, these cultures result in low daily output. Crabtree effect refers to the induction of fermentation, instead of respiration, in the presence of oxygen and high external glucose concentrations (Pfeiffer and Morley, 2014). For this reason, the Crabtree effect is an important phenomenon in aerobically grown yeasts, although it has disadvantages such as growth and cell yield if used in a manufacturing process. On the other hand, some yeast species do not show the crabtree effect, and consequently, they can achieve higher cell densities in aerobic culture than Crabtree positive yeasts (Imura et al., 2020). Examples of Crabtree positive yeast species can be listed as *S. cerevisiae*, *T. glabrata*, *Schizosaccharomyces pombe*, and *B. intermedius*. Furthermore, *C. utilis*, *H. nonfermentans*, *K. marxianus*, *P. stipites*, *Kluyveromyces*, and *P. pastoris* are the Crabtree negative yeast species (Urk et al., 1990; Hagman et al., 2013).

An additional classification of the yeast species can be done as methylotrophic and non-methylotrophic yeasts. Methylotrophic yeasts, such as *Candida*, *Pichia*, *Hansenula*, and *Torulopsis* (Gellissen et al., 1992), can utilize methanol as a sole carbon and energy source. Adaptation to growth on methanol is associated with induction of methanol oxidase, *MOX* (also referred to as alcohol oxidase, *AOX*), dihydroxyacetone synthase *DAS*, and several other enzymes involved in methanol metabolism (Sreekrishna and Kropp, 1996). On the other hand, non-methylotrophic yeasts do not have the ability to use methanol as a carbon source, but they can utilize a wide range of carbon sources for energy production. Furthermore, *S.cerevisiae*, *K.lactis*, *Y.lipolytica*, and *P.stipitis* are the yeasts that belong to this category (Fernández et al., 2016).

Among all yeasts, *S. cerevisiae*, the first eukaryote whose genome was fully sequenced, is one of the most preferred species for the expression of a wide variety of recombinant proteins. As well as common advantages of yeasts, *S. cerevisiae* has some further advantages for application of it in bioprocesses such as it is designated as generally regarded as safe (GRAS) strain and its well-characterized organism with the high-throughput data collection. Furthermore, it has tolerance for environmental

stress and low oxygen levels, and high glycosylation capacity (Çelik and Çalık, 2012).

Another yeast species commonly involved in recombinant protein production as a host organism is *Pichia pastoris*, also named *Komagataella phaffii*. Some additional advantages, such as tightly regulated and efficient promoters and a strong tendency for respiratory growth as opposed to fermentative growth, bring this species one step ahead of other yeasts for protein expression (Cregg et al., 2009).

Table 2. 2 *Properties of most commonly used species in bioprocesses (Gustavsson, 2018) (Mattanovich* et al., *2012)*

Expression System	***Pichia pastoris***	***Saccharomyces cerevisiae***	***Escherichia Coli***	**Chinese hamster ovary**
Growth media	Inexpensive, Minimal	Inexpensive, Minimal	Inexpensive, Minimal	Expensive, Complex
Cell growth	Rapid	Rapid	Rapid	Slow
Secretion	Secretion to extracellular environment	Secretion to extracellular environment	Secretion to extracellular environment	Secretion to periplasm
Easy to Scaling up	yes	yes	yes	no
Common Product	Insulin	Bioethanol	Human growth hormone	Monoclonal antibodies
Cost	Cost-effective	Cost-effective	Cost-effective	Expensive

2.1.1 *Pichia pastoris*

In recent years, the yeast *Pichia pastoris*, which is a methylotrophic yeast that can be genetically engineered to express proteins for basic research and industrial use, has become one of the most successful and popular host systems for heterologous protein production with several significant advantages (Çalık et al., 2015; Cereghino et al., 2002). In 1920, promising yeast for bioprocess application, *Pichia pastoris*, firstly isolated from the exudate of a chestnut tree in France, and A. Guilliermond characterized it as Zygosaccharomyces pastori (Guilliermond, 1920). The name "*Pichia pastoris*" was given to this species by H. Phaff, who isolated the further strains of it from trees in California in 1956 (Phaff et al., 1956). Afterward, in the 1970s, Philips Petroleum introduced *Pichia pastoris* for commercial production of single-cell proteins (SCP) as high-protein animal feed due to the ability to reach high cell densities. SCP production has become uneconomical since the oil crisis, which occurred in 1973, caused an increase in the price of methanol (Cereghino and Cregg, 2000). Whereas in the 1980s, *Pichia pastoris* developed as a heterologous protein expression system.

Three fundamental properties of *Pichia pastoris* made this species an ideal host for bioprocesses. The first one of these advantages is that it can be easily manipulated at the molecular genetic level. Secondly, it can express proteins at high levels, both intracellularly and extracellularly. Finally, it can perform many 'higher eukaryotic' post-translational modifications, such as glycosylation, disulfide bond formation, and proteolytic processing (Cereghino et al., 2002). Moreover, notably, the reengineering of the N-glycosylation pathway has enabled the production of heterologous proteins with human-like N-glycan structures (Mattanovich et al., 2009).

Another crucial characteristic of *P. pastoris*, which makes it an ideal host, is it is a methylotrophic yeast so that it can utilize methanol as a sole source of carbon and energy with the help of the alcohol oxidase I gene (AOX1) (Ogata et al., 1969). Moreover, *Pichia pastoris* has a very efficient and tightly regulated promoter, AOX1

promoter, which is derived from the AOX1 gene of the methanol utilization pathway. In particular, the use of this promoter makes high levels of foreign protein expression possible (Couderc and Baratti, 1980; Çelik and Çalık, 2012).

One another most essential properties of the *P. pastoris* is that it is described as a Creb-negative yeast (Mattanovich et al., 2009). In other words, *P. pastoris* prefers a respiratory rather than a fermentative mode of growth. Fermentative organisms produce by-products such as ethanol and acetic acid, and these compounds can reach toxic levels in a high cell density environment. (Cereghino et al., 2002). For this reason, with its respiratory mode of growth feature, *Pichia pastoris* becomes the reason for preference. Figure 2.2 exhibits the central carbon mechanism of *Pichia pastoris*.

Finally, its ability to grow on inexpensive and minimal medium and ability to secrete r-proteins to fermentation, which makes purification easier, are other vital factors that make *Pichia pastoris* an ideal host for recombinant protein production processes (Çalık et al., 2015)

Recently, reclassification of *P. pastoris* into a new genus, Komagataella, was made (Yamada et al., 1995). Indeed, this genus split into three species, K. pastoris, K. phaffii, and K. pseudopastoris (Kurtzman, 2005). Two of these new species, K. pastoris and K. phaffii, have strains that are used in biotechnological applications (Mattanovich et al., 2009). The strains GS115 and X-33 belong to K. phaffii species. On the other hand, although a name change has occurred in the literature, it is still known as *Pichia pastoris*.

All *P. pastoris* expression strains are derivatives of the wild type strain NRRL-Y 11430 and, the most commonly used strains are auxotrophic mutants (e.g., GS115) and protease-deficient strains (e.g., SMD1163, SMD1165, SMD1168) (Çelik and Çalık, 2012). Research Corporation Technologies (Tucson, AZ, USA) has the patent of expression system of *Pichia pastoris*, and the expression kit from Invitrogen Corporation (Carlsbad, CA, USA) can be used for research use (Çelik and Çalık, 2012).

Table 2. 3 *Advantages and disadvantages of P. pastoris expression system*

Advantages	**Disadvantages**
o High yield o High expression levels o High cell densities o Enabling easy purification o Rapid growth o Defined and inexpensive medium o Enabling easy scaling-up o Broad pH range o Performing PTMs o Genetically stable strains o Having strong promoters, P_{AOX1}, P_{GAP} o GRAS	o Long cultivation time o High proteolytic activity o Requiring bioreactor application for high-level production o Ability to perform non-native glycosylation and hypermannosylation

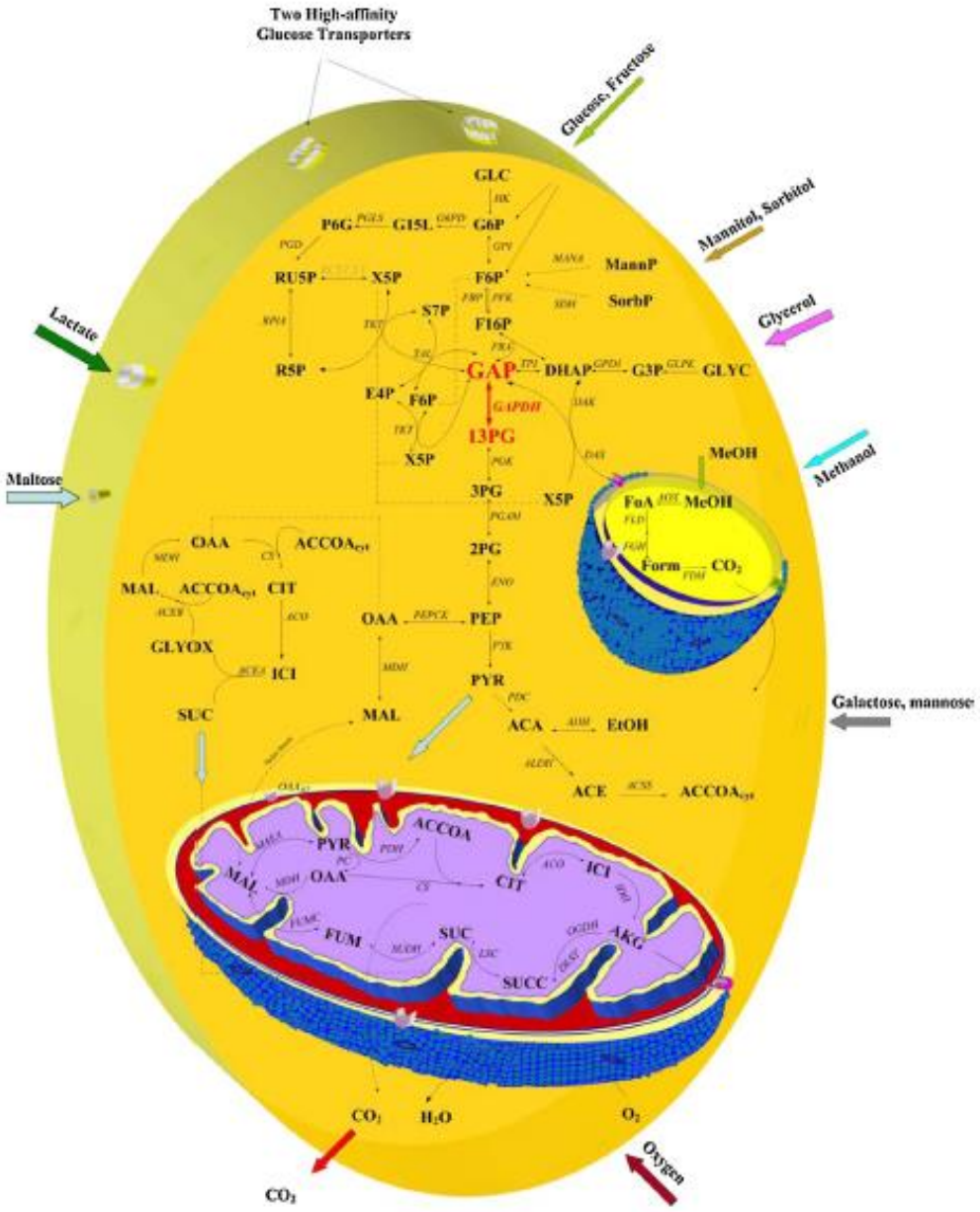


Figure 2. 2 Central carbon metabolism of *Pichia pastoris* (Çalık et al., 2015)

2.2 Promoters

Expression of the respective gene of interest (GOI) is essential for recombinant protein production. Furthermore, the expression of the GOI consists of several essential steps like transcription, translation, folding and possible post-translational modification, as well as the correct targeting by the heterologous host organism. In both prokaryotes and eukaryotes, the initial transcription of the GOI is often a key step in heterologous protein production. (Markides 1996 ;Porro et al., 2005 ;Desai et al.,2010 ;Vogl & Glieder, 2013) Transcription, conversion of a DNA sequence into an RNA transcript by RNA polymerase II (PolII), is the starting step of gene expression, and in transcription, all-protein-coding and many non-coding genes are transcribed. Transcription usually begins at the transcription start site (TSS), which is located at the 5th end of a gene, also known as the gene start. The TSS is located within a core promoter, which is a short sequence that spans 50 base pairs upstream and downstream of the TSS. The transcription machinery, which includes Pol II and its accompanying general transcription factors (GTFs), uses the core promoter as a binding platform. (Hampsey 1998; Haberle et al.,2018) In other words, a promoter is a DNA sequence enabling and regulating transcription initiation. (Ottoz and Rudolf, 2018) Consequently, strong and controlled promoters are vital tools for efficient expression systems because transcription, the initial stage in protein synthesis, is a critical step in r-protein creation. (Porro et al., 2005; Gasser et al., 2019) Promoters of yeast generally include two functionally and physically distinct regions, core promoter and upstream element. (Struhl, 1982; Beier et al., 1982; Guarente and Mason, 1983; Ottoz and Rudolf, 2018). The core promoter is the area that carries the bare minimum of information required to initiate transcription, regardless of whether or not it is regulated. (Beier et al., 1982; Faye et al., 1981) The core promoter serves as a recruiting platform for RNA polymerase II [21]. To bind the promoter and become competent for transcription initiation, RNA polymerase II requires the help of general transcription factors. (Venters et al., 2011) TATA-binding protein (TBP), a generic transcription factor, detects the TATA element, a

DNA sequence enriched for T and A, which is located at various distances upstream of the TSS(s). The interaction between TBP and the TATA region causes RNA polymerase II and other general transcription factors to be recruited to the core promoter in a stepwise manner. This causes the preinitiation complex (PIC) to form, which is required for transcription to begin. (Hahn et al., 1985; Zhang and Dietrich, 2005; Rhee and Pugh, 2012)

The upstream element confers regulation by recruiting transcription factors (Guarente, 1984). Transcription factors bind to the upstream element at transcription factor binding sites, which are distinct and well-known DNA patterns (TFBSs). TFBSs are both required and sufficient for a promoter to be regulated (Giniger et al., 1985). A schematic representation of the structure of a promoter and eukaryotic gene transcription is given in Figure 2.3.

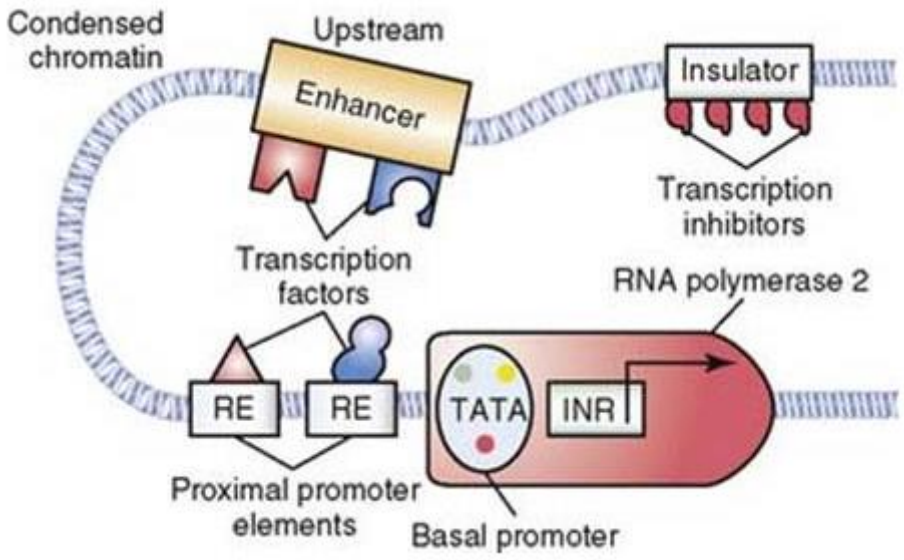


Figure 2. 3 Structure of a promoter and eukaryotic gene transcription

In the light of this information, it can be concluded that strong and controllable promoters are a crucial tool for efficient heterologous protein production. The powerful methanol inducible promoters present in *P. pastoris* and other methylotrophic yeasts are one of their most advantageous features. Certain methanol

utilization (MUT) pathway genes are totally repressed when grown on repressive carbon sources (e.g., glucose) but substantially activated when grown on methanol, accounting for up to 20%–30% of total intracellular protein. MUT promoters are attractive tools for heterologous protein production because they are tightly regulated at the transcriptional level. As a result, the majority of published information on the control of *P. pastoris* promoters focuses on the most often utilized traditional *P. pastoris* promoters, the promoters of the MUT pathway gene PAOX1, as well as the glycolysis gene PGAP. (Vogl and Glieder, 2013) The most often used promoters can be classified into two categories: constitutive and inducible promoters.

2.2.1 Constitutive Promoters

Constitutive promoters result in steady expression under a variety of circumstances. Strong constitutive promoters that drive high-level transcription are frequently utilized to obtain high levels of expression of important enzymes. (Xiong et al., 2018) In other words, a constitutive promoter has a generally steady activity that is unaffected by external stimuli. In most situations, the activity of constitutive promoters is linked to the rate of growth, which is determined by the amount of glucose, yeast's preferred carbon source. (Ottoz and Rudolf, 2018) One of the most common constitutive promoters of *P. pastoris*, Glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter, was characterized by Waterham *et al*. in 1997 as reaching as high expression level as P_{AOX1} promoter.

One of the most important features of the GAP promoter, which makes it preferable for non-toxic protein production, is that its constitutive expression shortens and simplifies the processing time because there is no need for induction. By this means, high protein titer and cell densities can be reached under the control of the GAP promoter. The ability to be used with other carbon sources than methanol is another critical feature of the GAP promoter. The GAP promoter can reach high cell densities and protein titers with glucose and glycerol, which are non-toxic for cells. This eliminates methanol usage and makes the GAP promoter a good alternative for

pAOX1. Besides the GAP promoter, there are other constitutive *P. pastoris* promoters that provide high expression levels that have been reported by researchers. For instance, two naturally occurring promoters around the pyruvate branch point, namely the pyruvate decarboxylase (PPDC) and pyruvate kinase (PPYK) promoters, as viable substitutes for constitutive PGAP, each with distinct expression patterns and strengths was reported in 2018 by Massahi and Çalık. In table 2.4, constitutive promoters of *P. pastoris* are given.

Table 2. 4 *Constitutive promoters of P. pastoris (Çalık* et al., *2015) (Massahi and Çalık, 2018)*

Gene Name	**Gene Product**
GAP	Glyceraldehyde 3-phosphate dehydrogenase
PGK1	Phosphoglycerate kinase
TEF1	Translation elongation factor 1 α
PDC	Pyruvate decarboxylase
PYK	Pyruvate kinase

2.2.2 Inducible Promoters

Environmental stimuli cause inducible promoters to vary their expression levels dramatically. To obtain high levels of expression of important enzymes, strong constitutive promoters that drive high-level transcription are frequently utilized. (Xiong et al., 2018) Alcohol oxidase (P_{AOX1}) is the most often used inducible promoter, which catalyzes the first step of the methanol utilization pathway, conversion of methanol to formaldehyde, of *P. pastoris*. (Ergün, Huccetogullari, Öztürk, Çelik, & Çalık, 2019). It is strongly induced on methanol; on the other hand, it is completely repressed under the growth source other than methanol. One other inducible promotor of *P. pastoris* is the alcohol dehydrogenase 2 (P_{ADH2}) promoter

which catalyzes the reaction that converts ethanol to acetaldehyde (Cregg & Tolstorukov, 2012). This promoter is induced on ethanol. Since it does not require methanol, which is toxic, it is a good alternative to recombinant protein production. In table 2.5, inducible promoters of *P. pastoris* are given.

Table 2. 5 *Constitutive promoters of P. pastoris (Çalık* et al., *2015)*

Gene Name	Regulation	Gene Product
AOX1	Induced on methanol	Alcohol oxidase 1
DAS	Induced on methanol	Dihydroxyacetone synthase
FLD1	Induced on methanol and methylamine	Formaldehyde dehydrogenase
PEX8	Induced on methanol or oleate	Peroxisomal matrix protein
ADH2	Induced on glycerol and ethanol, repressed by methanol and glucose	Alcohol dehydrogenase

2.2.3 Engineered Promoters

Naturally occurring promoters (NOPs) usually provide limited opportunities due to their complex regulation mechanisms and highly dependent working principles, which hamper their production efficiency and increase metabolic burden over the host cell. Synthetic biology extends on promoter engineering, aiming to increase the transcription rate through the generation of novel regulatory circuit(s) by engineering binding of the activating transcription factors (TFs). Promoter strength and regulation are cumulative effects of short and distinct DNA motifs that facilitate binding of cellular transcriptional machinery. To increase the transcription rate in yeasts, promoter engineering methods, such as random mutagenesis (Alper et al., 2005), chimeric promoter design (Blazeck et al., 2012), modification of transcription

factor binding sites (Ata et al., 2017), tuning of nucleosome architecture (Curran et al., 2014), synthetic core promoter design (Vogl et al., 2014), and hybrid-architectured engineered promoter variants (EPVs) design (Ergün et al., 2019a; Ergün et al., 2020) have been used.

Recently, two novel engineered promoter variants, which were also used in this study, were designed and constructed. The first one is the EPVs of P_{ADH2} which were designed by engineering of transcription factor binding sites (TFBSs) determined by *in silico* analyses and manual curation systematically by single-handed replacement of specified TFBSs with synthetic motifs for Mxr1, Cat8, and Aca1 binding to increase r-protein expression through the generation of novel regulatory circuits in *P. pastoris* over the EUT pathway. Compared with P*ADH2* at t = 20 h of batch-fermentations, the hybrid-architectured EPV $P_{ADH2\text{-}Cat8\text{-}L2}$ allowed the highest increase in enhanced r-protein expression as 4.8-fold on ethanol and 3.8-fold on methanol (Ergün et al., 2019a). The second one is the EPVs of P_{AOX1}, which was designed replacing specified *cis*-regulatory DNA elements with synthetic Adr1, Cat8, and Aca2 *cis*-acting DNA elements for Mxr1, Cat8, and Aca1 binding, respectively, and the expression over methanol-free substrate-utilization pathway(s) was transcriptionally rewired and was converted methanol-dependent *Pichia pastoris alcohol oxidase 1(AOX1)* promoter (P_{AOX1}) expression into a non-toxic carbon-source-regulated system. Compared with P_{AOX1} on methanol, the expression on ethanol is increased with PAOX1/Cat8-L3 (EPV used in this study) to 74% and increased with PAOX1/Adr1-L3/Cat8-L3 to 85% (Ergün et al., 2020). Demir and Çalık (2020) designed double double-promoter expression systems (DPESs) with the EPVs $P_{ADH2\text{-}Cat8\text{-}L2}$ and P_{mAOX1} to enhance and upregulate-deregulated gene expressions in *P. pastoris* on ethanol and compared the DPESs with the constituent single-promoter expression systems (SPESs). The transcription period and strength of each constituent in the DPESs were determined by expressing reporter red fluorescent protein and enhanced green fluorescent protein under PADH2-Cat8-L2 and PmAOX1, respectively. In batch-cultivations on ethanol, superior DPES performances were investigated in extracellular human growth hormone production

and compared to its constituent SPES constructed with PADH2-Cat8-L2. Furthermore, ethanol fed-batch bioreactor operations (FBBOs) were designed for enhanced recombinant human growth hormone (rhGH) production in *Pichia pastoris* constructed with novel hybrid-architectured *ADH2* promoter, P*ADH2-Cat8-L2*, and the parameters in the fed-batch fermentation of ethanol were investigated, and the boundaries of the production domain in terms of the design parameters were determined (Wehbe et al., 2020).

2.3 Target Proteins

Reporter proteins have tremendous importance in both bioprocesses and metabolic engineering applications since it enables the detection or measurement of gene expression. Furthermore, fluorescent bioimaging of single molecules, intact organelles, live cells, and whole organisms is an extensively used biochemistry, biotechnology, cell, and developmental biology approach. (Stepanenko et al.,2008) GFP was first isolated from the jellyfish *A. victoria* in the early 1960s by Osamu Shimomura. (Zimmer M. 2002) This invention led to a revolutionary impact on studies related to cell biology and almost every area of biomedical science since these "Molecular flashlights" have the advantage of being used as protein tags to track the subcellular localization of a protein of interest within the cell environment. Furthermore, they can also be used as genetically engineered fluorescent biosensors to monitor dynamic changes in signaling molecules in the context of living cells. For this reason, three distinguished researchers, Drs. Osamu Shimomura, Martin Chalfie, and Roger Tsien were awarded the 2008 Nobel Prize in Chemistry to discover and develop the green fluorescent protein *Aequorea victoria* (avGFP), the prototypical member of a diverse family of fluorescent and non-fluorescent proteins expressed by various marine organisms. (Sample et al., 2009)

Especially after cloning green fluorescent protein (GFP)1, fluorescent proteins (FPs) have become standard imaging tools for cell biologists. (Costantini *et al*. 2015) GFP, a complete genetically encoded tag, has become a unique tool that enables direct

visualization of structures and processes in living cells and organisms. Subsequently, intense interest in the structure, biochemistry, and biophysics of GFP-like fluorescent proteins emerged, resulting in an avalanche of scientific publications on FPs and their applications to solve fundamental molecular and cell biology problems. Nowadays, fluorescent proteins are being used in wide-variety applications (Figure 2.4), such as drug screening, protein interactions, and promotor activity, to screen the function and organization of living systems. (Chudakov et al., 2010)

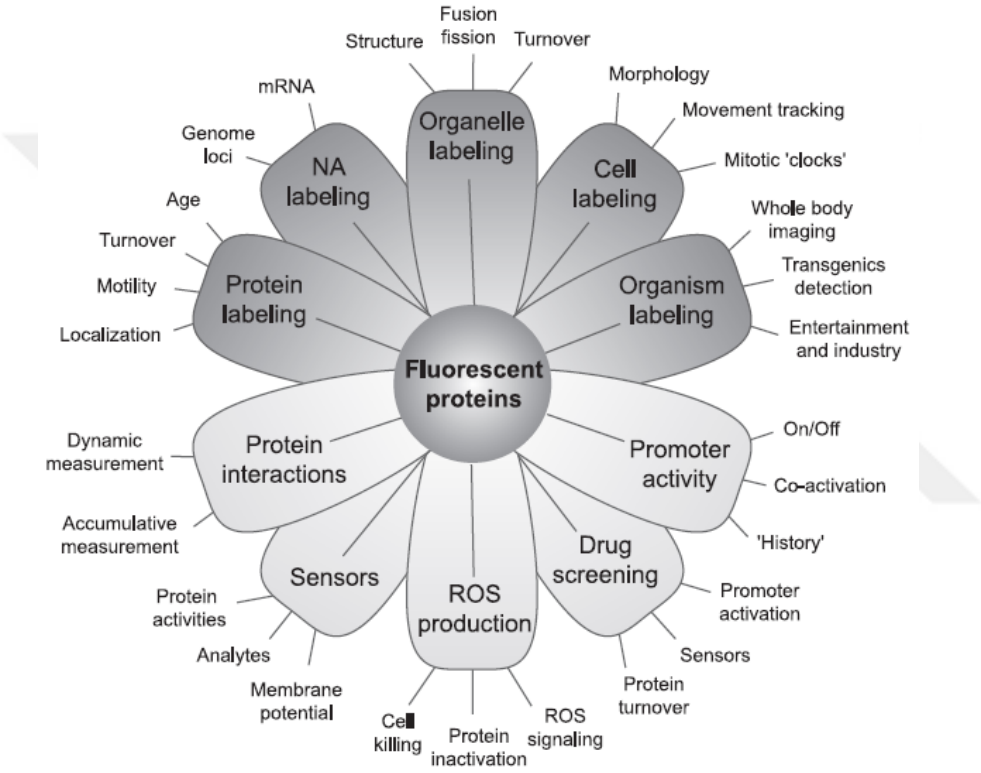


Figure 2. 4 Main areas of applications of fluorescent proteins (Chudakov et al., 2010)

Fluorescent proteins generally consist of 220-240 aminoacids residues (approximately around 25 kDa), which fold into a barrel formed by 11 β -sheets that accommodate a distorted internal helix, and all FPs have the ability of self-generation of the intrinsic chromophore from three amino acids at positions 65-67 without requiring co-factors or enzymatic components. Furthermore, FPs have a rigid β-

barrel shell, and it protects the chromophore from the environment and restricts its flexibility (Stepanenko et al.,2011). Besides these key features, fluorescent proteins also have a wide-diverse color scale covering almost the whole visible spectrum from violet and to far-red and this color diversity. This current palette, which has various colors of monomeric FPs, enables multicolor protein labeling experiments with as many as six colors and potentially more. As a result, it provides new possibilities for multiparameter imaging of structures and processes in living systems. (Kogure et al., 2006) The spectral diversity of available monomeric FPs is shown in Figure 2.5. (Chudakov DM et al., 2010)

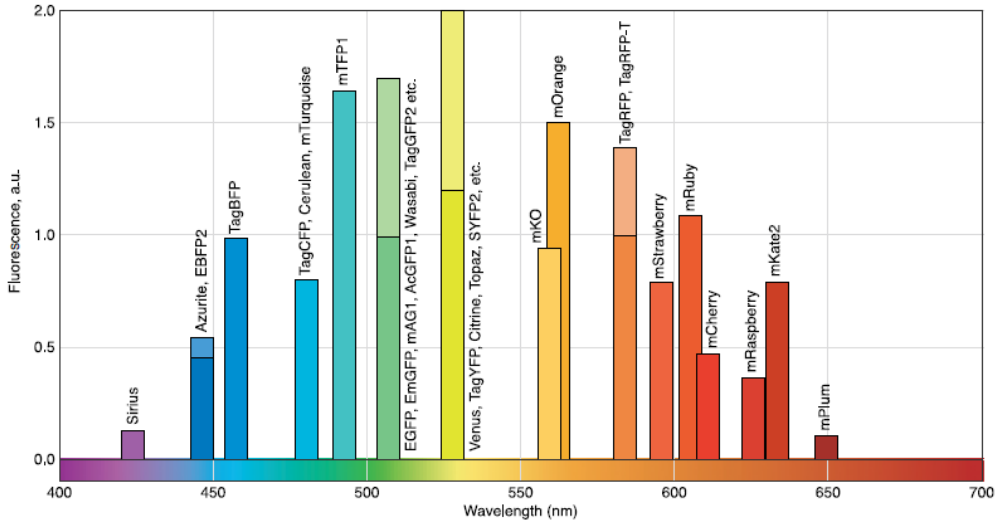


Figure 2. 5 Spectral diversity of available monomeric FPs.

Thus, there are several critical characteristics for the practical use and preference criteria of FPs for bioimaging. One of them is the brightness of FPs. The brightness of the fluorophore used has a significant impact on the sensitivity of any fluorescence detection technology. The high brightness of an FP provides additional benefits for fluorescent labeling in biological systems. The brighter FPs, in particular, require a smaller dosage of excitation light, resulting in less phototoxic consequences. (Chudakov et al., 2010) Another critical feature of FPs in use is the maturation rate. After protein folding and chromophore maturation, which includes many covalent modifications, FP fluorescence develops. Orange and red FPs have the most covalent

modifications. The rate-limiting phase for the FP to become fluorescent is usually chromophore maturation. Maturation can take anything from a few minutes (103, 304) to hours or even days, depending on the specific FP, oxygen content, and temperature. (Nagai et al., 2002; Evdokimov et al., 2006) The maturation rate should be rather quick for practical usage to get a high and stable fluorescent signal.
The maturation half-time of most FPs in use ranges from 40 minutes to 12 hours, which is long enough to label cells, organelles, and proteins of interest, as well as undertake numerous quantitative assays. However, FPs with very fast maturation is required for several applications, such as early detection of promoter activation, tagging proteins of interest with a brief lifespan, or monitoring single translational processes. Fast-maturing yellow FPs, for example, permitted real-time observation of the synthesis of a single protein molecule. (Yu et al., 2006) Finally, one of the cornerstones of a successful imaging experiment is FP photostability, which becomes a vital feature for extended time series, accumulating weak fluorescence signals, and quantitative measures, such as FRET approaches. The protein shell protects the chromophore of FPs from the environment. This is substantially responsible for wild-type FPs' generally strong photostability and low phototoxicity, which is likely vital for their native function in sea organisms. With the exception of KillerRed, most mutant forms of natural FPs have low phototoxicity. Successful mutant versions also maintain wild-type FPs' naturally excellent photostability. (Chudakov et al., 2010)

Considering the mentioned features and criteria, eGFP and mApple fluorescent proteins were preferred as reporter genes in this study.

2.3.1 *eGFP* Protein

GFP is a protein consisting of 238 amino acids and weighs 27 kDa. It gives an excitation peak at 395 nm and an emission peak at 509 nm. Other researchers have developed many variants of GFP. Heim and Tsien (1996) made the most remarkable improvement with the S65T point-mutation, which achieved an increase in

fluorescence amount and stability, and the excitation peak shifted to 488 nm. The F64L mutation enabled the generation of the enhanced GFP variant and made it possible for this protein to be used by mammalian cells (Cormack et al., 1996). In *P. pastoris*, GFP has been used for many different purposes (such as protein tracking, organelle tracking, marking, and screening of promoter variants) (Hartner et al., 2008; Qin et al., 2011; Ruth et al., 2010; Prielhofer et al., 2013; Heiss et al., 2015; Puxbaum et al., 2016; Rossanese et al., 1999; Sakai et al., 1998; Sjöblom et al., 2012).

2.3.2 *mApple* Protein

The discovery and development of Green Fluorescent proteins have led to the discovery and development of many new fluorescent proteins. Red Fluorescent Protein (KFP) is one of the fluorescent protein discoveries with a wide area of use in many research areas. KFP (or DsRed), first isolated from *Discosoma striata*, is a protein consisting of 225 amino acids and weighs approximately 25.9 kDa. (Matz et al., 1999) It gives the excitation peak at 558 nm and the emission peak at 583 nm. The main disadvantages limiting the use of DsRed are the lack of photostability and the long time between protein synthesis and fluorescent expression. To eliminate these disadvantages, different variants of KFP have been developed by many researchers. One of the most important developments was the monomeric red fluorescent protein (mRFP1), with random and directed mutations, thus shortening the maturation time of the fluorescent protein by ten times. In addition, with this development, the excitation peak of mKFP1 shifted to 584 nm and the emission peak to 607 nm, providing spectral separation from other fluorescent proteins (Campbell et al., 2002). Subsequently, a number of mKFP1 variants (mFruits) were generated by directed mutation, resulting in a significant improvement in stability and fluorescence (Shaner et al., 2008). One of the variants with the highest fluorescence amount and stability among the mKFP1 variants is the mApple fluorescent protein. This protein weighs approximately 27.3 kDa and gives excitation and emission peaks

at 568 and 592 nm, respectively. Just like YFP, RFP is also found in *P. pastoris*; It has been used for many different purposes such as protein tracking, organelle track, marker, and designing new plasmid variants (Raschmanova et al., 2019; Barrero et al., 2018; Zheng et al., 2019; Chen et al., 2012; Schroder et al., 2008; Chen et al., 2017).

2.4 Ethanol Utilization Pathway of *Pichia pastoris*

With the discovery of promoters and genes that can be induced in the presence of ethanol, ethanol has become a prominent choice as a carbon source in recombinant protein production. In addition, the fact that the use of methanol has a toxic effect, especially in the production of pharmaceutical products, and makes purification difficult, makes different carbon sources interesting. In addition to these, the fact that ethanol is more cost-effective and easy to purify carbon sources than glucose is one of its prominent features.

Ethanol, the fermentation pathway's final product, is also utilized as a carbon source by the cells. It is usually used by Adh isoenzymes once the preferred carbon source has been depleted. (De Smidt, Du Preez, & Albertyn, 2008). In *P. pastoris*' metabolism, utilization of ethanol starts with oxidation of ethanol to acetaldehyde.

Adh enzymes catalyze both the reduction of acetaldehyde to ethanol and the oxidation of ethanol to acetaldehyde, making them important actors in the fermentation and ethanol use pathways; furthermore, experimental investigations revealed that *ADH2* is the only gene responsible for the consumption of ethanol in *P. pastoris* (Karaoglan et al., 2016; Karaoglan et al., 2019). After oxidation of ethanol to acetaldehyde, by the activity of acetaldehyde dehydrogenase enzyme, produced ethanol converted to acetate. Aldehyde dehydrogenases are involved in the conversion of acetaldehyde to acetyl-CoA during non-fermentable carbon sources development, as well as the breakdown of hazardous aldehydes accumulated under stress situations (Aranda and Olmo, 2003). The cytoplasmic Ald6 and mitochondrial

Ald4 and Ald5 enzymes oxidize acetaldehyde to acetic acid; the first two employs NADP as a co-factor, while the latter uses NAD (Boubekeur et al., 2001; Saint-Prix et al., 2004; Simpson-Lavy and Kurpeic, 2019). Then, two acetyl-CoA synthetases: Acs1 and Acs2, conjugate acetate to coenzyme-A in the cytoplasm (Kratzer & Schuller., 1995; Van den Berg & Steensma., 1995; Simpson-Lavy and Kurpeic, 2019). ACS enzyme is responsible for the conversion of acetic acid to acetyl-CoA in yeast. Chen and co-workers (2013) reported that over-expression of *ACS1* or *ACS2* in *S. cerevisiae* increased intracellular acetyl-CoA levels two to five-fold (Chen *et al*. 2013). Two genes in yeast have been found to code for proteins involved in acetyl-CoA synthetase activity. *ACS1* encodes a mitochondrial isoform that is present during respiratory growth and glucose repressed (van den Berg et al. 1996), whereas *ACS2* encodes a cytosolic isoform that is necessary for anaerobic glucose growth (van den Berg and Steensma 1995; Heit et al. 2018). In *S. cerevisiae*, acetyl-CoA synthetase is encoded by two genes, expression of *ACS1* is repressed by glucose and derepressed by ethanol or acetate; on the other hand, ACS2 is essential for growth on glucose and is expressed constitutively (Shiba et al., 2007). After conjugation acetate to coenzyme-A,-finally, produced acetyl-CoA enters the TCA cycle. In figure 2.6, a diagram of the ethanol utilization pathway of the *P. pastoris and* TCA cycle is given.

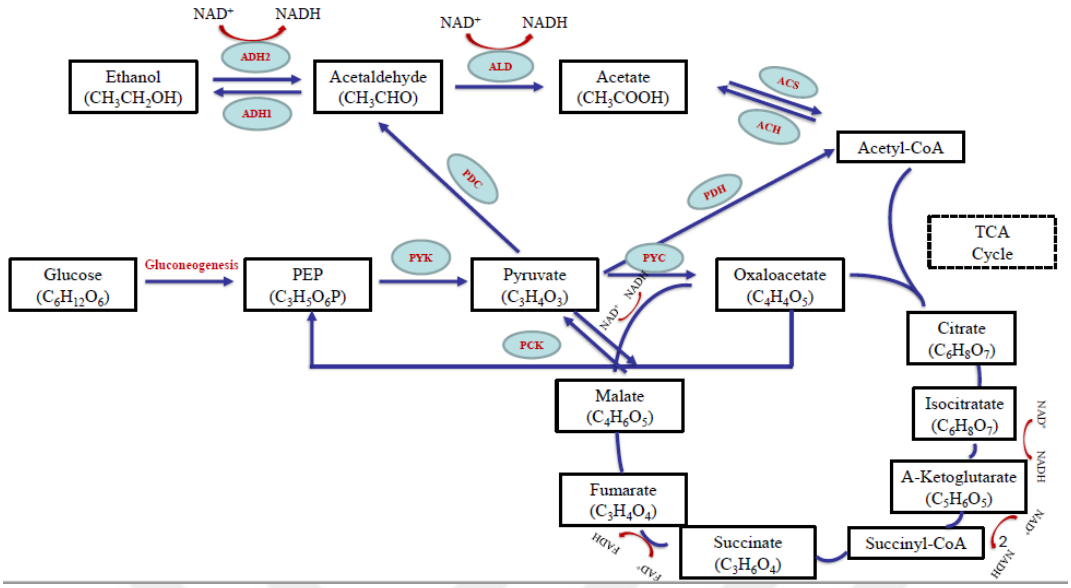


Figure 2. 6 A diagram of ethanol utilization pathway and TCA cycle in *Pichia pastoris*

2.5 Metabolic Engineering

Metabolic engineering is the technique of enhancing a cell's production of a specific substance by improving its genetic and regulatory systems. These are chemical networks that use a series of biochemical reactions and enzymes to let cells transform basic materials into molecules required for cell survival. Metabolic engineering aims to quantitatively describe these networks, compute the yield of valuable products, and pinpoint the sections of the network that limit product output. (Yang et al., 1998) The metabolic pathways of numerous species can be modified using recombinant DNA technology. Bacteria, fungi, plants, and animal cells are all used as host organisms, and the properties of interested pathways are used to guide selection. Amino acids, biofuels, secondary metabolites, recombinant proteins, and polymers are all critical industrial products. Engineers use methodologies including rational metabolic engineering, evolutionary engineering, inverse metabolic engineering, secretory pathway engineering, and process optimization to boost productivity. The goal to reduce costs, boost yields, increase efficiency, increase customer acceptance, and employ renewable feedstocks to provide a more ecologically friendly process

than traditional chemical techniques drives metabolic engineering applications. (Wuest et al., 2011) In short, the ultimate goal of metabolic engineering can be defined as using these organisms to produce valuable substances on an industrial scale in a cost-effective manner. (Milne et al., 2020)

Several metabolic engineering strategies have been applied for years to increase the production capacity of the cells. The first one is *Engineering biosynthetic pathways.* Successful examples of the application of this strategy generally include

- Amplification of the gene coding the biosynthetic pathway's rate-limiting enzyme removes the bottleneck and enhances r-protein production (Cremer et al., 1991; Yang, et al., 2007). Schematic representation of this approach was given in figure 2.7.

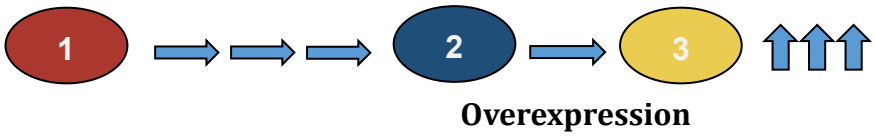


Figure 2. 7 Strategy of metabolic engineering for the production of a desired chemical: overexpression of therate-limiting enzyme.

- Inhibiting competing metabolic pathways that use the same substrate is another technique to overproduce a specific pathway. As a result, the substrate is metabolically directed towards the target molecule. (Kulkarni R., 2016) Schematic representation of this approach was given in figure 2.8.

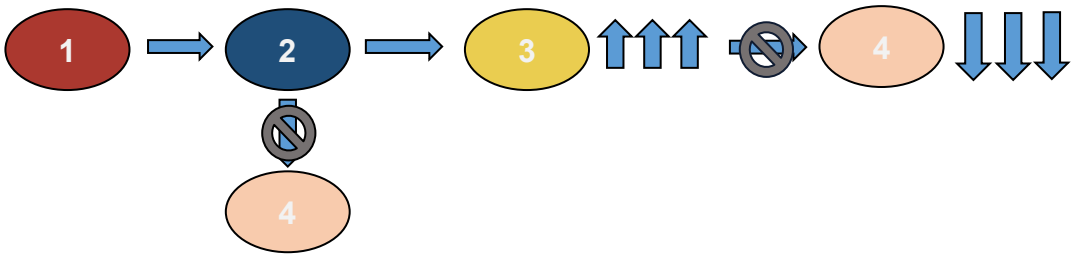


Figure 2. 8 Strategy of metabolic engineering for the production of a desired chemical: inhibition of the competing pathway.

- Amplification of the gene coding for the branch-point enzyme to redirect metabolic flow from the common intermediate to another amino acid. (Katsumata R. et al., 1986 ;Yang, S.T et al., 2007)
- Introducing heterologous enzymes with alternative control structures allows them to skip a regulatory step in the biosynthetic pathway (e.g., feedback inhibition or repression). (Guillouet et al., 1999; Yang et al., 2007)

  The second common approach is engineering central metabolism. When constraints in the terminal routes are overcome, central metabolism provides precursors and energy, which may limit amino acid production in overproducing bacteria.

  Engineering the central metabolism is difficult since it is governed by a complex and global regulatory system that is poorly understood.

  Another common metabolic engineering approach is transport engineering. With this approach, Mutant strains with altered transport systems can maintain a low intracellular level of the product amino acid and hence are not subject to feedback control. Overproduction of the amino acid is possible in mutants with an active efflux mechanism and defective amino acid absorption. Without completely deregulating the necessary biosynthetic pathways, mutants with an active efflux system and impaired amino acid absorption can overproduce the amino acid. (Ikeda M., 2003; Yang et al., 2007)

To conclude, the most commonly used metabolic engineering approaches include deletion or overexpression of genes in biosynthetic pathways to increase productivity. In table 2.6 some common metabolic engineering approaches were summarized. Furthermore, in this study, overexpressing a gene encoding the rate-limiting enzyme of the biosynthetic pathway was applied as a metabolic engineering approach.

Table 2. 6 *Some common metabolic engineering strategies*

Strategies for increasing productivity and/or yield
Amplify the gene (cluster) coding for the rate-limiting enzyme to eliminate the bottleneck
Amplify the gene coding for the branch-point enzyme to redirect the metabolic flow
Delete key genes in the branched pathways to eliminate non-productive reactions
Modulate enzyme expression levels to avoid the accumulation of intermediates
Introduce genes/enzymes with different control architectures to avoid feedback inhibition or repression
Increase the supply of precursor
Change cofactor dependency of key enzymes to balance redox and/or improve energy and growth efficiency
Engineer the transport systems to improve product efflux and/or substrate uptake

2.6 Transcript-level Expression Analysis

Comparative transcriptome analysis is a powerful tool for the study of global dynamics of yeast gene expression in response to various environmental, genetics, and chemicals cues. In recent years hundreds of transcriptome studies performed either with microarray or RNAseq platforms provided an enormous resource for understanding the whole-genome regulatory networks and mechanisms of yeast gene expression in response to various stresses (Taymaz-Nikerel et al., 2016). Enzyme activity and metabolite abundances determine cellular metabolic fluxes. Biochemical methods disclose the effect of certain substrates or regulators on enzyme kinetics, but they don't account for how metabolite and enzyme concentrations vary between physiological states and hence how biological activities are regulated (Heckett et al.,

2016). Despite the extensive knowledge of metabolic reaction networks that resulted, the means by which metabolic reaction rates (fluxes) are controlled remain incompletely understood, even in highly studied model microbes. Therefore, considering transcriptomics and metabolites together has become the right strategy to investigate the effect of metabolic engineering on the intracellular reaction network of the organism.

By analyzing the differential transcriptome of recombinant strains with 12 copies and a single copy of phospholipase A_2 gene (*PLA* $_2$) from *Streptomyces violaceoruber,* the global effect of overexpression of the recombinant protein, which often leads to a severe burden on the physiology of *K. phaffii* and triggers the cellular stress was investigated, and it was revealed that the 12-copy strain suffered heavy cellular stress. The genes involved in protein processing and stress response were significantly upregulated due to the burden of protein folding, and secretion, three most upregulated heat shock response genes (*CPR6*, *FES1*, and *STI1*) were co-overexpressed in *K. phaffii* and proved their positive effect on the secretion of reporter enzymes (PLA_2 and prolyl endopeptidase) by increasing the production up to 1.41-fold, providing novel helper factors for rational engineering of *K. phaffii* (Yu et al., 2017). Tredwell et al. (2017) showed by finding robust correlations between metabolites and *KAR2*/*PDI1* transcript levels that, monitoring the unfolded protein response (UPR) levels by gene transcription is certainly feasible at a potentially high-throughput miniaturized scale and can therefore address a key bottleneck in the pathway to produced secreted proteins. Furthermore, the relative mRNA levels of KAR2 and PDI genes were estimated using qRT-PCR to investigate whether reduced porcine insulin precursor, PIP, expression in high copy strains could be the result of the bottleneck at the secretory pathway, and results showed that the mRNA ratio of KAR2 increased with the enhanced gene dosage, suggesting that the unfolded protein response was more pronounced in high copy strains, while no significant differences were observed based on the data of mRNA ratio of PDI gene (Zhu et al., 2009).

2.7 Specific Ethanol Consumption Rates in Air-filtered Shake-flask Bioreactors

Specific ethanol uptake rate is a key parameter for understanding and comparing how the modifications on the ethanol pathway affect cell metabolism and growth. In batch operations, fermentation typically begins at t=0 h after inoculation and continues throughout the process without the use of a feeding cell or carbon source. Thus, mass conservation equations for the ethanol in batch processes can be derived as follows (Çalık et al., 2016);

$$r_{EtOH} = \frac{d(C_{ETOH}.V)}{dt} \tag{2.1}$$

Where r_{EtOH} is the substrate consumption rate (g/L h), V is the volume of the air-filtered shake-flask bioreactor (L). r_{EtOH} can be defined with a first-order kinetic equation;

$$-r_{EtOH} = q_{EtOH}.C_X \tag{2.2}$$

where q_{EtOH} is the specific ethanol uptake rate (g/g h), and C_X is the cell concentration in the shake-flask bioreactor (g/L). By the insertion of equation 2.2 into 2.1, the following equation can be derived;

$$-q_{EtOH}.C_X.V = \frac{dC_{EtOH}}{dt}.V + \frac{dV}{dt}.C_{EtOH} \tag{2.3}$$

Since the volume of the shake-flask is constant, dV/dt equals zero, and by dividing both sides of the equation by V, equation 2.2 becomes;

$$-q_{EtOH}.C_X.= \frac{dC_{EtOH}}{dt} \tag{2.4}$$

Finally, the specific ethanol uptake rate equation can be obtained as;

$$q_{EtOH} = -\frac{1}{C_X}\left(\frac{dC_{EtOH}}{dt}\right) \tag{2.5}$$

CHAPTER 3

MATERIALS AND METHODS

3.1 Kits and Enzymes, Chemicals and DNA Ladders

Thermoscientific (Thermo Fisher, USA) is the supplier of all the DNA ladders and kits which, are used in PCR purification, plasmid isolation, and gel extraction. Thermoscientific and NEB are the suppliers of all the restriction enzymes used in this study also. All chemicals and solutions used in this work were obtained from Sigma-Aldrich and Merck Millipore.

3.2 Strains, Plasmids, Primers, and Maintenance

E. coli DH5α (Invitrogen, USA) cells were used to amplify the constructed plasmid and also used in the cloning experiments. *P. pastoris* X-33 (Invitrogen, Carlsbad, Ca, USA) strain was used as the expression strain for cloning experiments, which include the construction of mApple producing strains and overexpression of Alcohol dehydrogenase enzyme, ADH2. To be able to construct predesigned plasmids, pGAPZα A (Invitrogen) plasmid was used as parent plasmid. All microorganism strains were kept in 25% glycerol stock solution at -80° for long-term storage. Oligomer (Ankara) synthesized primer stocks as 100µM, and primer stocks were diluted to 10µM to satisfy working concentration in this study. All primer stocks, diluted primers, plasmid DNA were kept at -20°C for storage. The list of used and constructed strains and plasmids in this study and designed primers for the construction of recombinant plasmids are given in Table 3.1 and Table 3.2, respectively.

Table 3. 1 *Constructed strains and plasmids*

Microorganism	Plasmid	Comment	Source
***E. coli* DH5α**	pGAPZα-A	Backbone plasmid	The plasmid was synthesized and stored in *E. coli* DH5α
***E. coli* DH5α**	pmAOX1::*eGFP*	The plasmid carrying P*AOX1/Adr1-L3/Cat8-L3* with *eGFP*	The plasmid was synthesized and stored in *E. coli* DH5α (Ergün et al., 2020)
***E. coli* DH5α**	pADH2-Cat8-L2::*mApple*	Plasmid PADH2-Cat8-L2 with *mApple* (selection marker: NTC)	The plasmid was synthesized and stored in *E. coli*
***E. coli* DH5α**	pADH2-Cat8-L2::*hGH*	The plasmid carrying PADH2-Cat8-L2 with *hGH* (selection marker: NTC) with its native secretion signal	The plasmid was synthesized and stored in *E. coli* DH5α (Demir and Çalık, 2020)
***E. coli* DH5α**	pPM2eno-*Hph*	The plasmid carrying the hygromycin resistance gene	Gift from Mattanovich's lab

Table 3.1 (Continued)

***E. coli* DH5α**	pADH2-wt::*ADH2*	The plasmid carrying P_{ADH2wt} with ADH2 (selection marker: Zeocin)	Constructed in this study
***E. coli* DH5α**	pADH2-wt::*ADH2*	The plasmid carrying P_{ADH2wt} with ADH2 (selection marker: Hygromycin)	Constructed in this study
***E. coli* DH5α**	pmAOX1::*mApple*	The plasmid carrying P*AOX1/Adr1-L3/Cat8-L3 with mApple*	Constructed in this study
***P. pastoris* X-33**	pmAOX1::mApple	*P. pastoris* strain used for screening P*AOX1/Adr1-L3/Cat8-L3* (single copy)	Constructed in this study
***P. pastoris* X-33**	pmAOX1::mApple + pGAP:: *eGFP*	*P. pastoris* strain used for screening P*AOX1/Adr1-L3/Cat8-L3* and P*GAP*(parent plasmid) at the same time	Constructed in this study

Table 3.1 (Continued)

***P. pastoris* X-33**	pmAOX1::mApple + pADH2-Cat8-L2::*mApple*	*P. pastoris* strain used for screening P*AOX1/Adr1-L3/Cat8-L3* and pADH2-Cat8-L2::*mApple* (parent plasmid) at the same time	Constructed in this study
***P. pastoris* X-33**	pADH2-wt::*ADH2*	*P. pastoris* strain which overexpresses ADH2	Constructed in this study

Table 3. 2 *Designed primers that were used in this study.*

Primer	Sequence (5'-3')	Restriction Enzyme	Used for
Forward_P$_{ADH2\text{-}wt}$	GTC*GGATCCCTGCAG*TCCTTTTTAC	*Bam*HI and *Pst*I	Amplification of promoter and enzyme complex
Reverse_ADH2	GAT*GGTACC*TTATTTGGAAGTGTCCACAACG	*Kpn*I	Amplification of promoter and enzyme complex
Forward_P$_{mAOX}$	GAA*GGCGCGCC*AACATCCAAAGACGAAAG	*Asc*I	Amplification of promoter variants
Reverse_P$_{mAOX}$	CGCCCTTGCTCACCATCGTTTCGAATAATTAGTTGTTTTTTGATCTTCTC		Amplification of promoter variants
Forward_mApple	GAGAAGATCAAAAAACAACTAATTATTCGAAACGATGGTGAGCAAGGGCG		Amplification of mApple
Reverse_mApple	GCC*TCTAGA*TTACTTGTACAGCTCGTCCATGCC	*Xba*I	Amplification of mApple

Table 3.2 (Continued)

Forward_Hyg	CAA*GGATCC*AGCT CAGGGGCATGATG TG	*Bam*HI	Amplification of Hygromycin
Reverse_Hyg	GCG*ACATGT*GCACTT AACTTCGCATCTGGG	*Pci*I	Amplification of Hygromycin
Forward_P*AOX1-Cat8-L3*	GGCGGATCCAGATCT AACATCCAAAGAC	*Bam*HI	Amplification of promoter variant
Reversep_P*AOX1-Cat8-L3*	GAAGTAAGTGTTCGT TATCTAATGGCATCG TTTCGAATAATTAGT TG		Amplification of promoter variants
Forward_ACS1	CAACTAATTATTCGA AACGATGCCATTAG ATAACGAACACTTAC TTC		Amplification of *ACS1*
Reverse_ACS1	GTTGGTACCTCATTT GCGGGCATC	*Kpn*I	Amplification of *ACS1*

3.3 Genetic Engineering Methods

3.3.1 Plasmid Isolation

Plasmid isolation from *E. coli* cells was executed by using the GeneJET Plasmid Miniprep Kit (ThermoFisher, USA) in line with the manufacturer's guidance. As the prerequisite of the plasmid isolation, the overnight growth of cells in LB medium at 37°C and 200 rpm with proper antibiotics, Zeocin, Nourseothricin (NTC), or Hygromycin, was provided. Grown cell culture was harvested by centrifugation at 6800g for 5 minutes, and pellets were used in plasmid isolation by applying the following protocol:

1. Resuspension of the cells with 250µl Resuspension solution and transfer the mixture into the 1.5ml microcentrifuge tube.
2. Addition of 250µl of Lysis solution and mixing completely
3. Immediate addition of 350µl Neutralization solution and mixing
4. Centrifugation for 5 minutes after complete mixing
5. Washing and elution steps
6. Storage of pure DNA plasmid at -20°C

3.3.2 Genomic DNA Isolation

Genomic DNA isolation from *P. pastoris* cells was carried out by using the Wizard Genomic DNA purification Kit (Promega, USA)) in line with the manufacturer's instructions. Before genomic DNA isolation, the growth of the cells in YPD medium at 30°C and 200 rpm with proper antibiotics, Zeocin, Nourseothricin (NTC), or Hygromycin, was provided. 1 ml of the grown cell culture was harvested by centrifugation at 14000g for 2 minutes, and pellets were used in genomic DNA isolation by applying the manufacturer's protocol. Isolated DNA kept at -20°C for further usage.

3.3.3 Agarose Gel Electrophoresis

For control and visualization of the PCR products and isolated plasmids, and extraction of digested DNA fragments, agarose gel electrophoresis was carried out. Visualization of the DNA fragments was accomplished based on the length of the fragments with agarose gel electrophoresis (AGE). AGE was conducted by using the Mini-sub® Cell GT Cell system (Bio-Rad, CA, USA). The following description is the general protocol of the AGE.

1. Each usage requires dilution of the 50X TAE buffer to 1X buffer.
2. Achievement of the desired resolution requires the dissolution of the proper amount of agarose, depending on the length of the DNA fragment, in the 1X TAE buffer. Dissolution of the 8g/L and 12g/L agarose is appropriate for the visualization of DNA fragments and the gel extraction, respectively.
3. Heating the solution to the boiling point and boiling until the mixture becomes clear
4. Addition of EtBr, which gives the final concentration of 0.4 μg/mL (1.75 μL EtBr for 50 ml 1X TAE) following the cooling of the solution
5. Fitting the combs into the plastic trays to form wells
6. Pouring the gel into the plastic trays and waiting for approximately 30 minutes to firm the gel
7. Usage of 1 μL of the 6X DNA gel-loading dye for a 1-5 μL DNA sample. In some cases, the sample requires dilution with sterilized water
8. Mixing of the DNA solution by pipetting and loading into the wells
9. Loading of 5 μL DNA marker (Thermoscientific), which based on the length of the DNA fragment that visualized, into the well
10. Attachment of the lid to the gel-electrophoresis tank and connection of the electrical leads
11. Performing the gel-electrophoresis at 90V for 30-80 minutes, based on length of the DNA fragment and gel concentration

12. Visualization of the DNA bands under UV-light with ethidium bromide filter Hamamatsu Digital CCD Camera was used for visualization and saving the images.

3.3.4 Extraction of DNA Fragments

Agarose Gel Electrophoresis was performed on the DNA samples to separate desired DNA fragments. Extraction of the DNA fragments from the gel was provided by using a sterile razor blade under UV light, and extracted fragment imbed to Eppendorf tube. The extraction of the fragment by razor procedure should be applied quickly and carefully in order not to hazard the DNA. Purification of a single DNA fragment was performed using the gel extraction kit (Thermo Scientific, MA, USA).

3.3.5 Gel Elution

GeneJet Gel Extraction Kit (ThermoFisher, USA) was used for gel elution by applying the manufacturer's instructions. The following description is the general protocol of gel elution.

1. Extracted agarose gel with DNA fragment, which is less than 400 mg, is put in 1.5 mL Eppendorf tubes.
2. Based on the weight of the interrupted agarose gel, the binding buffer was added at the ratio of 1:1.
3. The incubation of the Eppendorf tube at 65°C for 10 minutes provided for dissolving the gel.
4. Transfer of the dissolved gel into the GeneJet purification column is carried out, and the column is centrifuged at 14000g for 1 minute.
5. Discard of the fluid which passed from the column

6. Centrifugation at 14000g 1 more minute provided after de addition of 700 µL Wash buffer
7. Since ethanol may be hazardous for the DNA, to prevent the presence of the ethanol in the column, centrifugation of the empty column at 14000g for 2 minutes was provided.
8. To be able to dispose of all the ethanol, the column put in a new 1.5 ml of Eppendorf tube.
9. Addition of 25-50µl of ultrapure water or elution buffer into the column and centrifugation at 14000g for 1 minute provided.
10. Storage of the purified DNA at the -20°C

Room temperature is the requirement of all centrifugation steps.

3.3.6 PCR Purification

GeneJET PCR Purification Kit (ThermoFisher, USA) was used by following the manufacturer's protocol for purification of the PCR products.

3.3.7 DNA Sequencing

Verification of the sequences of the inserted and amplified DNA was carried out in SenteBiolab (Ankara).

3.3.8 Transformation to *E. coli*

Transformation of the constructed plasmids, given in Table 3.1, into the *E. coli* DH5α was carried out. For this purpose, competent *E. coli* cells were prepared by using the $CaCl_2$ method (Sambrook and Russell, 2001), whose procedure is described below. Firstly, *E. coli* DH5α cell micro-bank stock maintained at -80 ° C was transferred to LB Agar medium (15 g / L agar, 5 g / L NaCl, 10 g / L tryptone, 5 g / L yeast extract). After 24 hours of incubation, a single colony was selected and transferred to LB broth (10 g / L tryptone, 5 g / L NaCl, 5 g / L Yeast extract). When the concentration of the cells, which were grown in an orbital-agitated batch bioreactor at 37 ° C, 200 rpm at, was reached to the OD600 = 0.35-0.40, they were taken a 50 ml falcon and set on the ice for 10 minutes. Then, centrifugation at 2700g for 10 minutes at 4°C was provided to be able to obtain pellets. Afterward, the addition of 30 ml of 80 mM MgCl 2-20 mM CaCl 2 solution to the cell pellet was carried out, and the cells were dissolved in the solution by swirling and centrifuged at 2700 g for 10 minutes at 4 ° C. After discard of the supernatant, the pellet resuspended with the 2ml of 0.1M CaCl2 solution. In this way, competent cells were constituted. 50 µL of the competent cells were used for each transformation. The addition of 50 ng of DNA in a maximum of 10µl volume into the competent cell solution was carried out, and the mixture was kept on ice for 30 minutes. Subsequently, cells were subjected to heat shock at 42 ° C for 90 seconds in the water bath and set onto the ice for 5 minutes. Following that, the addition of 900 µL of LB liquid medium into the mixture was carried out. Cells were amplified in an orbital-agitated bioreactor at 37 ° C, 200 rpm for 60 min. Grown cells were inoculated into selective LB Agar medium with proper antibiotics, NTC (50µl/100ml), or Hygromycin (150µl/100ml). After 16-20 hours of incubation at 37°C, at least ten colonies, which proliferated on the LB agar, were selected for further verification. For verification, firstly, colony PCR was performed. Afterward, selection of at least four clones of confirmed strains was provided, and plasmid isolation by Plasmid MiniPrep Kit (Thermoscientific) according to the manufacturer's instructions was carried out. Polymerase chain reaction (PCR) was

performed to amplify the desired DNA from isolated plasmids. By the use of agarose gel electrophoresis, the length of the DNA fragments was verified. Finally, for the formal verification, sequence analysis on the DNA fragments was carried out by Sentebiolab (Ankara). Microbank stocks were prepared and stored at -80°C for further usage. Plasmids from these cells were obtained by using the MiniPrep Plasmid Isolation Kit (Thermoscientific) in line with the manufacturer's protocol and kept at -20°C.

3.3.9 Transformation to *P. pastoris*

The lithium chloride method with linearized plasmid was followed for the transformation into wild-type *P. pastoris* cells. Firstly, wild-type *P. pastoris* x-33 cells were inoculated onto a YPD agar medium and incubated at 30°C for 48 hours. After incubation was completed, a single colony was chosen and inoculated into the YPD medium and incubated at 30°C and 200 rpm for approximately 14-16 hours until the concentration of the cells reaches OD_{600} of 0.8-1.0. Afterward, centrifugation at 4000g for 5 minutes at room temperature was provided to harvest the cells and obtain the pellets. The pellets were washed with 25 ml of sterile water and centrifugated at 1500g for 10 minutes at room temperature. Following that, the supernatant was discarded and resuspension of the cell pellet with 1ml of 100mM filtered sterilized LiCl solution was carried out. Then the solution was transferred into a 1.5ml Eppendorf tube and centrifugated at maximum speed for 15 seconds. Again, the supernatant was removed, and resuspension of the pellet with 400 µl of 100 mM LiCl solution was carried out. After this step, competent cells formed, and for each transformation, 50 µl of competent cells used—50 µl competent cell solution centrifugated at the maximum speed for 15 seconds and LiCl discarded by pipette. Then, the addition of the following components in the given order was carried out very slowly.

1. 240 µl of 50% PEG,
2. 36 µl 1 M LiCl
3. 25 µl 2 mg/ml single-stranded DNA
4. 5-10 µg of linearized plasmid DNA in a maximum of 50 µl of purified water

The tube was vortexed vigorously for approximately 1 minute, provide complete mixing, and incubated at 30°C without shaking for 30 minutes. Afterward, the tube was subjected to heat shock at 42°C in the water bath for 25 minutes and centrifugated at 6000-8000rpm for 15 seconds. Supernatant discarded, and the pellet was resuspended in 1ml of YPD medium and incubated at 30°C with shaking. Inoculation of 50,100,150 and 200µl of the medium onto the YPD agar containing appropriate antibiotics was carried out after 3 hours of incubation, and YPD agar plates were incubated at 30°C for 2-3 days. The putative transfectants were selected, and verification of the integration of the transferred gene was carried out by applying the colony PCR method, which follows 5 minutes of initial cell denaturation at 95°C. Among the colonies, which gave positive results from the colony PCR experiment, meaning that which carries the gene of interest, eight clones for each strain were selected and inoculated into the YPD media containing proper selective marker and incubated at 30°C and 200 rpm for 12-16 hours. Then, 25% of glycerol stocks were prepared and kept at -80°C for further investigations.

3.4 Construction of Strains and Plasmids

3.4.1 Construction of Recombinant Plasmids pADH2-wt::*ADH2*

3.4.1.1 Primer Design

For the amplification of the designed DNA sequences, Polymerase chain reaction (PCR) was used, and to be able to use PCR efficiently; primer design is one of the most crucial steps. In the primer design process, made benefit of Oligo analyzer 3.1 for some critical issues such as melting points (Tm), ΔG values, hetero, and

homodimer possibilities. The supplier of the designed primers was Sentegen (Ankara, Turkey). For the finding restriction sites, made benefit of SnapGene Viewer and Restriction Mapper. For the amplification of the target gene and insertion to the pGAPZαA vector, designed primers and restriction sites were used.

3.4.1.2 Amplification of the desired DNA

The usage of Polymerase chain reaction provided amplification of the desired DNA fragments. For the PCR experiments, special primers were designed and used. Optimization of some parameters such as extension time and annealing temperature were made to increase the efficiency of the PCR, and PCR experiments were carried out in thermocycler (Techne®, Flexigene, and TC-3000X). Amplification of the cassette, which contains the wild-type alcohol dehydrogenase promoter and alcohol dehydrogenase gene, P_{ADH2wt}::ADH2, was carried out from the genome of the *P. pastoris* wild type x-33 cell. PCR operation conditions for amplification of P_{ADH2}-*ADH2* and composition of PCR reaction mixture for P_{ADH2}-*ADH2* were demonstrated in Table 3.3 and 3.4, respectively.

Table 3. 3 *PCR operation conditions for amplification of* P_{ADH2}*-ADH2*

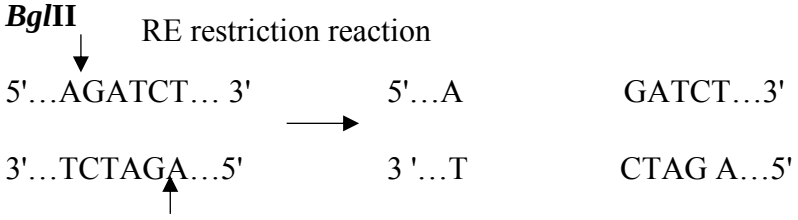
Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	65.7 °C	30 sec	
Extension	72 °C	80 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

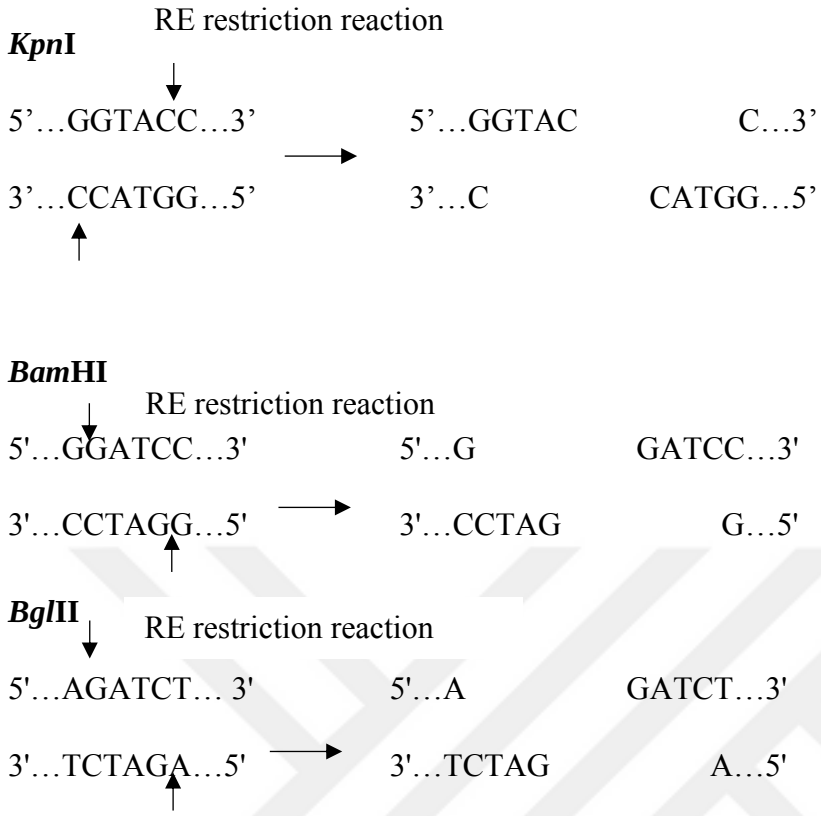
Table 3. 4 *Composition of PCR reaction mixture for P_{ADH2}-ADH2*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template DNA	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

3.4.1.3 Purification of the PCR Products and Digestion Reactions

To eliminate the interferants, PCR purification, conducted by using PCR purification Kit (ThermoScientific), was performed after the amplification of the desired DNA sequence by the usage of PCR. Designing the constructed plasmid containing pADH2:: *ADH2* cassette requires the restriction of both vector plasmid, pGAPZαA, and the insert, pADH2:: *ADH2*. For this purpose, proper restriction enzymes, which do not cut the desired gene, were selected. For the restriction of the vector plasmid (pGAPZαA), *Bgl*II, and *Kpn*I and the restriction of the insert, PADH2:: *ADH2*, *Bam*HI and *Kpn*I were selected.





Since the usage of the proper buffer optimizes the pH of the reaction medium, the addition of the appropriate buffer to the reaction mixture is required. *Kpn*I and *Bgl*II are incompatible with each other; these two restriction enzymes can not work in the same buffer. Thus, instead of double digestion, sequential digestion was applied with *Kpn*I and *Bgl*II to vector plasmid. Firstly, single digestion of the vector plasmid, pGAPZαA, was carried out by the usage of *Kpn*I. After the first digestion was completed, PCR purification was performed; afterward, the second digestion with the *Bgl*II was carried out. Both single digestions were carried out at 37°C for overnight, considering the complete digestion.

On the other hand, since *Bam*HI and *Kpn*I are compatible with each other, double digestion was applied for the digestion of the insert, pADH2:: *ADH2*. The double digestion procedure was carried out at 37°C for overnight. Conditions of single

digestion reaction with *Kpn*I and *Bgl*II reactions for pGAPZαA were tabulated in Table 3.5 and Table 3.6, respectively. Double digestion conditions with *Bam*HI and *Kpn*I for insert are demonstrated in Table 3.7.

Table 3. 5 *Conditions of single digestion reaction with KpnI for pGAPZαA*

Component	**20 µL Reaction**
10X *Kpn*I Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 60ng/µl)	16 µl
*Kpn*I	2 µl
H_2O	Up to 20 µl

Table 3. 6 *Conditions of single digestion reaction with BglII for pGAPZαA*

Component	**20 µL Reaction**
10X 0 Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 200ng/µl)	5 µl
*Kpn*I	2 µl
H_2O	Up to 20 µl

Table 3. 7 *Conditions of double digestion reaction with BamHI and KpnI for the insert*

Component	**20 µL Reaction**
10X 0 Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 60ng/µl)	16.7 µl
*Bam*HI	2 µl
*Kpn*I (2 fold of *Bam*HI)	4 µl
H_2O	Up to 20 µl

Inactivation of the digestion processes was provided at 80°C for 20 minutes in the water bath. After sequential digestion of the vector and double digestion of the DNA fragment, agarose gel electrophoresis was applied to the vector and DNA fragment for the verification of the successful double digestion. Desired parts of both vector and the insert extracted from the gel, and gel elution was carried out by the usage of the GeneJet Gel Extraction Kit (ThermoFisher, USA) in line with the manufacturer's protocol as explained in section 3.4.5.

3.4.1.4 Ligation reaction of the vector and the insert

The ligation of vector, pGAPZαA, and the insert, pADH2:: *ADH2*, was done by using the T4 DNA Ligase enzyme. The amount of the sufficient ligation mixture is designated as 20 µl. The ligation process was carried out at 16°C overnight, and the composition of the related ligation mixtures was tabulated in Table 3.8.

The optimum ratio of the vector to insert DNA and quantity of the vector are determined as 3::1 and 50 ng, respectively. Moreover, the quantity of the insert DNA was calculated by the application of the following equation:

$$Insert\ amount\ (ng) = 50\ ng\ vector \times \frac{Length\ of\ insert\ (bp)}{Length\ of\ vector\ (bp)} \times molar\ ratio \quad (3.1)$$

Inactivation of the T4 DNA ligase enzyme was accomplished by putting the mixture in the water bath for the heat shock at 65°C for 10 minutes, and the ligation mixture kept at -20°C for *E. coli* transformation.

Table 3. 8 *Composition of the ligation reaction mixture of pGAPZαA and ADH2-ADH2*

Component	**20 µL Reaction**
T4 DNA Ligation Buffer	2 µl
Vector DNA, pGAPZαA (DNA concentration is 25 ng/µl)	2 µl
Insert DNA, PADH2:: *ADH2* (DNA concentration is 10 ng/µl)	13.5 µl
T4 DNA Ligase	1 µl
H_2O	Up to 20 µl

After ligation, the construction of the desired plasmid was achieved and showed in Figure 3.1.

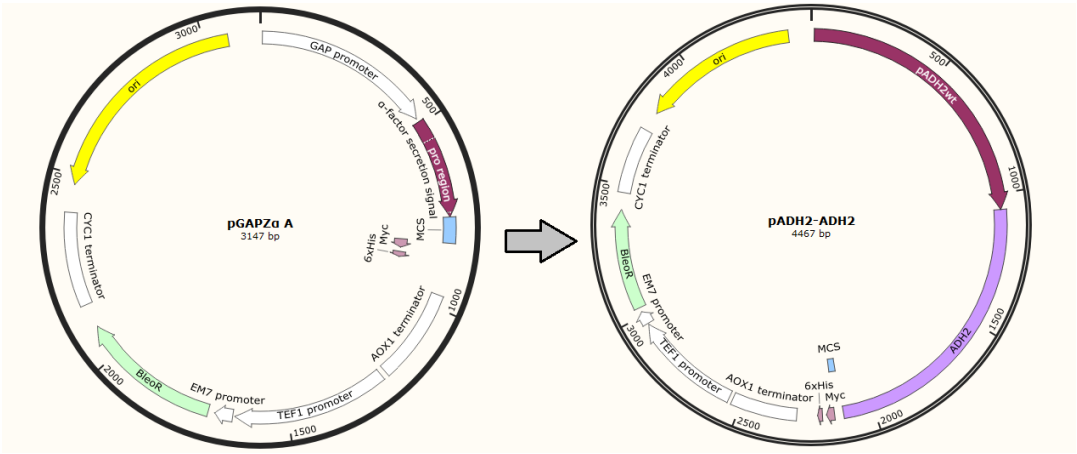


Figure 3. 1 Map of the constructed plasmid (pADH2:: *ADH2*) from pGAPZα

After ligation, a transformation of the designed plasmid to wild-type *E. coli* DH-5α cell was performed by the application of the LiCl method. 7.5 µl of the ligation mixture used for the transformation. Inoculation of transformed cells onto the LB agar medium containing the 25µg/ml zeocin, as a selective marker, was carried out and incubated for 16-18 hours. After incubation, selected colonies were inoculated onto the fresh LB agar medium, and the verification procedure was conducted with the help of colony PCR. Following the colony PCR, at least four colonies, which give the positive results for the verification of inserted plasmid, were selected, and plasmid isolation was carried out by the usage of Plasmid Miniprep Kit (Thermofisher). The sequence of the inserted DNA fragment was verified by gene sequencing.

Further studies for increasing the efficiency of the TCA cycle and metabolic flux may require the combination of different plasmids with this designed plasmid. For this reason, a new selection marker other than zeocin, which pGAPZα contains commercially, was essential. For this purpose, replacement of the antibiotic resistance gene of pADH2:: *ADH2* with Hygromycin with TEF promoter and TEF terminator was carried out.

The Hygromycin resistance gene, a gift from Mattanovich's lab, was amplified by the application of PCR by Q5 High Fidelity DNA Polymerase (Thermoscientific). PCR operation conditions for amplification of Hygromycin and the PCR reaction mixture composition were listed in Table 3.9 and Table 3.10, respectively.

Table 3. 9 *PCR operation conditions for amplification of hygromycin*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	68.8 °C	30 sec	
Extension	72 °C	55 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 10 *Composition of PCR reaction mixture for hygromycin*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template DNA	0.3 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

After amplification of the hygromycin resistance gene, vector plasmid, pADH2:: *ADH2*, and insert DNA, hygromycin resistance gene, both were double digested with the restriction enzymes, *Bam*HI and *Pci*I.

***Bam*HI**

RE restriction reaction

5'…GGATCC…3' 5'…G GATCC…3'

3'…CCTAGG…5' → 3'…CCTAG G…5'

***Pci*I**

RE restriction reaction

5'…ACATGT…3' 5'…A CATGT…3'

3'…TGTACA…5' → 3'…TGTAC A…5'

Double digestion process took place in the water bath at 37°C for 4.5 hours of vector plasmid and 10-12 hours for the insert. Composition of double digestion reaction mixtures of vector plasmid and insert tabulated in Table 3.11 and Table 3.12.

Table 3. 11 *Conditions of double digestion reaction with BamHI and PciI for vector plasmid*

Component	**20 µL Reaction**
10X Tango Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 110ng/µl)	9 µl
*Bam*HI	1 µl
*Kpn*I	1 µl
H_2O	Up to 20 µl

Table 3. 12 *Conditions of double digestion reaction with BamHI and PciI for the insert*

Component	**20 µL Reaction**
10X 0 Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 120ng/µl)	8 µl
*Bam*HI	1 µl
*Kpn*I	1 µl
H_2O	Up to 20 µl

Afterward of the double digestion, agarose gel electrophoresis was conducted for both insert and vector plasmid and desired DNA fragments extracted from the gel and purified by the usage of the GeneJet Gel Extraction Kit (ThermoFisher, USA) following the manufacturer's instructions. Then, ligation of the vector plasmid, pADH2:: *ADH2*, and the insert, hygromycin resistance gene, was carried by the ligation reaction with the T4 DNA ligase enzyme. The optimum ratio of the vector to insert DNA and quantity of the vector were designated as 3::1 and 50 ng, respectively. Moreover, the quantity of the insert DNA was calculated from Equation 3.1. Composition of the ligation reaction mixture of pADH2:: *ADH2* and Hygromycin given in Table 3.13. The ligation procedure was carried out at 16°C for 16 hours.

Table 3. 13 *Composition of the ligation reaction mixture of pADH2:: ADH2 and Hygromycin*

Component	**20 µL Reaction**
T4 DNA Ligation Buffer	2 µl
Vector DNA, pGAPZαA (DNA concentration is 20 ng/µl)	2.5 µl
Insert DNA, PADH2:: *ADH2* (DNA concentration is 20 ng/µl)	3.4 µl
T4 DNA Ligase	1 µl
H_2O	Up to 20 µl

After ligation, the same transformation procedure for *E. coli* transformation, which was explained in the previous part of this section, was repeated, and construction of the plasmid, which carries both pADH2:: *ADH2* cassette and hygromycin resistance gene completed. The designed plasmid was shown in Figure 3.2.

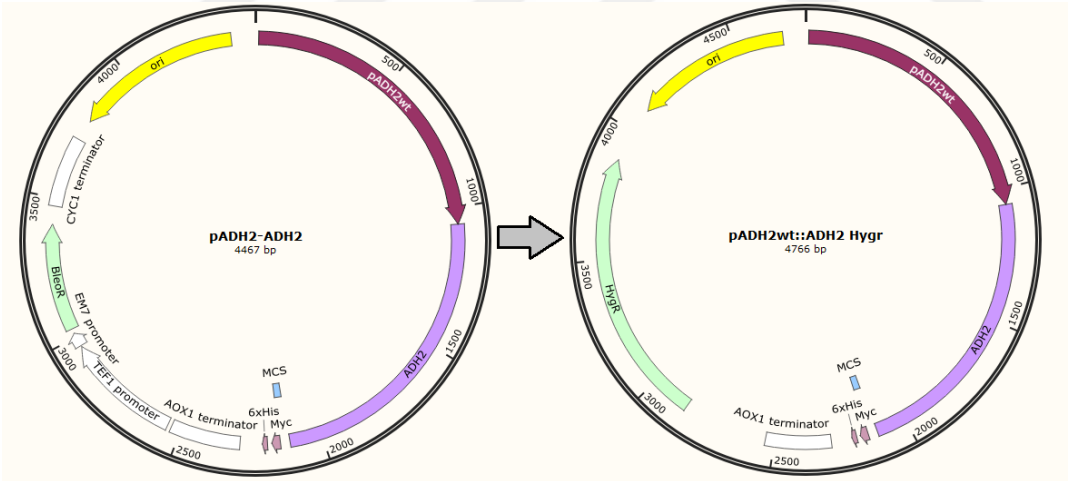


Figure 3. 2 Map of the constructed plasmid (pADH2:: *ADH2*) with a hygromycin resistance

3.4.2 Construction of Recombinant Plasmids pAOX1-Cat8-L3::*ACS1*

The following is the technique for constructing the plasmid pAOX1-Cat8-L3::*ACS1* mApple began with primer design. Sentegen provided related primers that were designed for use in the amplification procedures of desired genes (Ankara, Turkey).

3.4.2.1 Amplification of the desired DNA

The targeted DNA fragments were amplified using the polymerase chain reaction (PCR). Based on chosen primers, annealing temperatures and extension times were optimized to maximize PCR efficiency, and PCR reactions were performed in a thermocycler (Techne®, Flexigene, and TC-3000X). The pAOX1-v::eGFP plasmid and the wild type *P. pastoris* X-33 strain, respectively, were used to amplify the modified alcohol oxidase 1 promoter variant and the *ACS1* gene. Tables 3.14 and 3.15 show the PCR conditions for amplification of pAOX-Cat8-L3 and *ACS1*. In addition, Table 3.16 shows the composition of PCR reaction mixtures for pAOX1-v-*ACS1*.

Table 3. 14 *PCR operation conditions for amplification of pAOX1-Cat8-L3*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	64.3 °C	30 sec	
Extension	72 °C	35 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 15 *PCR operation conditions for amplification of ACS1*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	64.3 °C	70 sec	
Extension	72 °C	80 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 16 *Composition of PCR reaction mixtures for pAOX1-Cat8-L3 and ACS1*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template DNA	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

Following the amplification of the pmAOX1 and mApple fragments, the undesired PCR elements and byproducts were removed using the GeneJET PCR Purification Kit (ThermoFisher, USA). The fusing of the PAOX1-v fragment to the ACS1 enzyme was then performed using fusion PCR. Figure 3.3 shows the building logic algorithm for the pAOXCat8-L3::*ACS1* fragment.

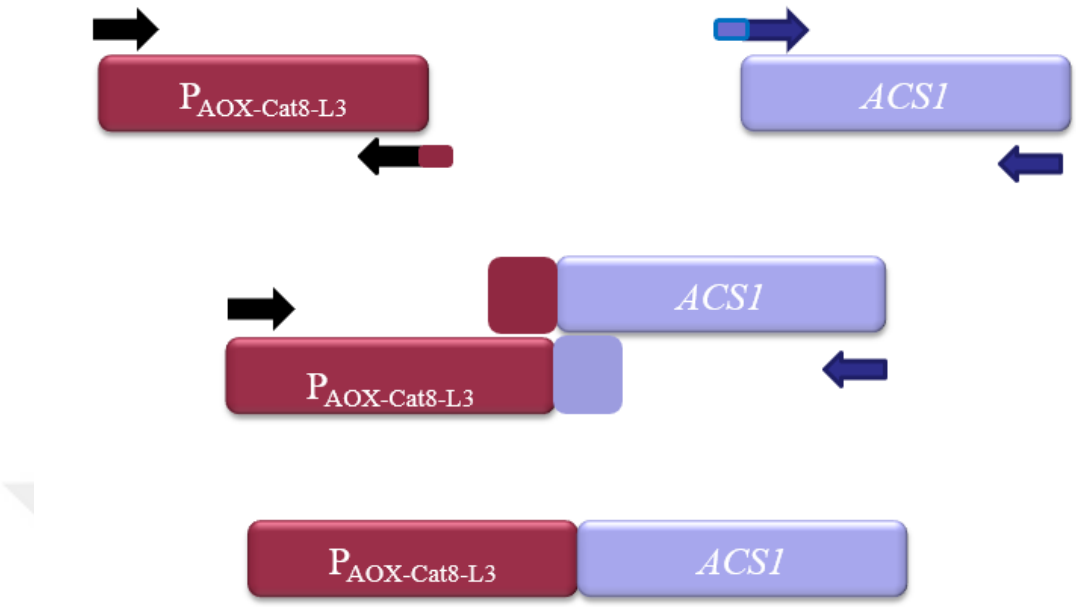


Figure 3. 3 Algorithm of construction logic of $P_{AOX1\text{-}Cat8\text{-}L3}$::*ACS1* fragment

Amplification of both pAOX1-v and *ACS1* fragments was achieved using forward and reverse pAOX1-v and *ACS1* primers, respectively. To be able to fuse these two fragments, the reverse primer of pAOX1-v and the forward primer of *ACS1* were created in such a way that their sequences overlapped. These two fragments have merged in a second PCR application, completing the production of the pAOX1-Cat8-L3::*ACS1* fragment. Tables 3.17 and 3.18 show the thermo-cyclic PCR settings and reaction components for overlap extension.

Table 3. 17 *Thermo-cyclic PCR operation conditions for construction of $P_{AOX1\text{-}Cat8\text{-}L3}$-ACS1*

Steps	Temperature	Time	Cycle
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	64.3 °C	30 sec	
Extension	72 °C	105 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

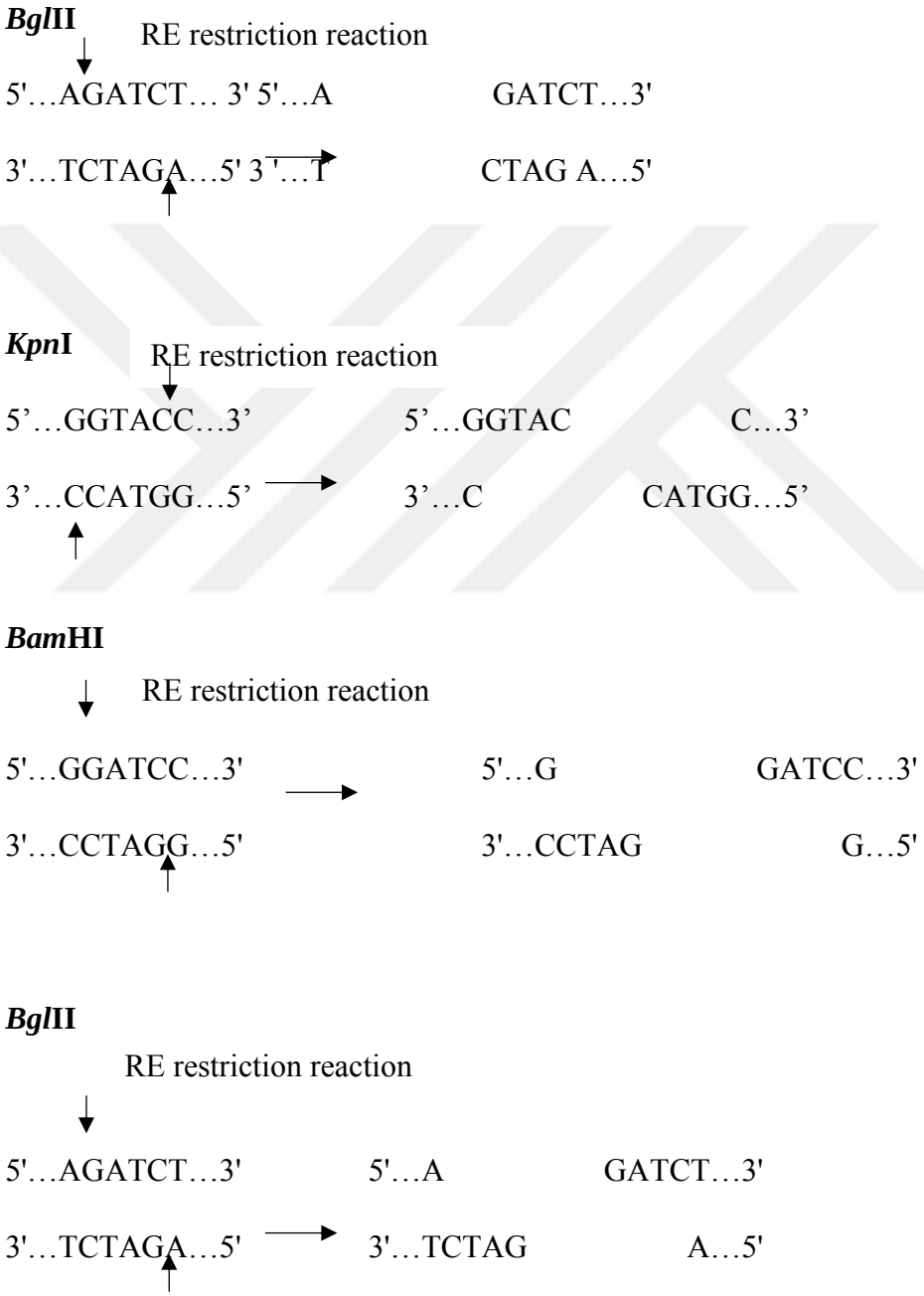
Table 3. 18 *Composition of PCR reaction mixture for construction of $P_{AOX1\text{-}Cat8\text{-}L3}$-ACS1*

Component	50 µL Reaction
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template 1 (mAOX1)	2 µL
Template 2 (mApple)	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

3.4.2.2 Purification of the PCR Products and Digestion Reactions

Agarose gel electrophoresis was done on SOE product, pAOX1-Cat8-L3-mApple, to remove meddling elements from the PCR reaction, then purified by gel elution with GeneJet Gel Extraction Kit (ThermoFisher, USA) according to the directions in

section 3.4.5. To produce the desired plasmid, both vector DNA, pGAPZα-A (KanMX), and insert DNA, pAOX1C8-L3-::eGFP must be digested with the identical restriction enzymes that were chosen in the primer design method. As a result, restriction enzymes *Bgl*II and *Kpn*I were used to digest plasmid, *Bam*HI, and *Kpn*I to digest vector plasmid, respectively.



The addition of the appropriate buffer to the reaction mixture is essential because the proper buffer optimizes the pH of the reaction medium. *Kpn*I and *Bgl*II are incompatible with each other; therefore, they cannot work together in the same buffer. Instead of double digestion, the vector plasmid was subjected to sequential digestion with *Kpn*I and *Bgl*II. To begin, *Kpn*I was used to perform single digestion of the vector plasmid, pGAPZα-A. After completing the first digestion, PCR purification was carried out, followed by second digestion with the *Bgl*II enzyme. Both single digestions were carried out overnight at 37°C to ensure complete digestion.

On the other hand, because *Bam*HI and *Kpn*I are compatible, the insert, pAOX-1Cat8-L3::*ACS1*, was double digested. The double digestion technique was carried out overnight at 37°C. Table 3.19 and Table 3.20 show the conditions of a single digestion process with *Kpn*I and *Bgl*II for pGAPZα-A, respectively. Table 3.21 shows the insert digestion conditions using *Bam*HI and *Kpn*I.

Table 3. 19 *Conditions of single digestion reaction with KpnI for pGAPZαA*

Component	**20 µL Reaction**
10X *Kpn*I Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 60ng/µl)	10 µl
*Kpn*I	1.8 µl
H_2O	Up to 20 µl

Table 3. 20 *Conditions of single digestion reaction with BglII for pGAPZαA*

Component	**20 µL Reaction**
10X 0 Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 200ng/µl)	5 µl
*Kpn*I	2 µl
H_2O	Up to 20 µl

Table 3. 21 *Conditions of double digestion reaction with BamHI and KpnI for the insert*

Component	**20 µL Reaction**
10X 0 Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 60ng/µl)	4.5 µl
*Bam*HI	2 µl
*Kpn*I (2 fold of *Bam*HI)	4 µl
H_2O	Up to 20 µl

The digestion processes were inactivated in the water bath at 80°C for 20 minutes. After sequential digestion of the vector and double digestion of the DNA fragment, the vector and DNA fragment was subjected to agarose gel electrophoresis to ensure that the double digestion was successful. The GeneJet Gel Extraction Kit (ThermoFisher, USA) was used to extract the desired sections of both the vector and the insert from the gel, and gel elution was performed according to the manufacturer's protocol, as stated in section 3.4.5.

3.4.2.3 Ligation reaction of the vector and the insert

The T4 DNA ligase enzyme was utilized to bind the vector and insert it together. The optimum amount of reaction mixture for the ligation reaction was determined to be 20 µl. The ligation procedure was completed at 16°C for 16 hours, and the ligation mixture composition is listed in Table 3.22. The optimal vector to insert DNA ratio and vector quantity were determined to be 3:1 and 50 ng, respectively. Equation 3.1 was also used to calculate the amount of insert DNA.

Table 3. 22 *Composition of the ligation reaction mixture of vector and insert*

Component	**20 µL Reaction**
T4 DNA Ligation Buffer	2 µl
Vector DNA, pGAPzα-A (DNA concentration is 50 ng/µl)	1 µl
Insert DNA, pAOX1-Cat8-L3:: *ACS1* (DNA concentration is 7 ng/µl)	16 µl
T4 DNA Ligase	1 µl

Afterward, ligation, the *E. coli* transformation procedure described in the previous section of this chapter, was carried out. The construction of the intended plasmid was completed with the ligation of the insert and vector and transformation to *E. coli*, as shown in Figure 3.4.

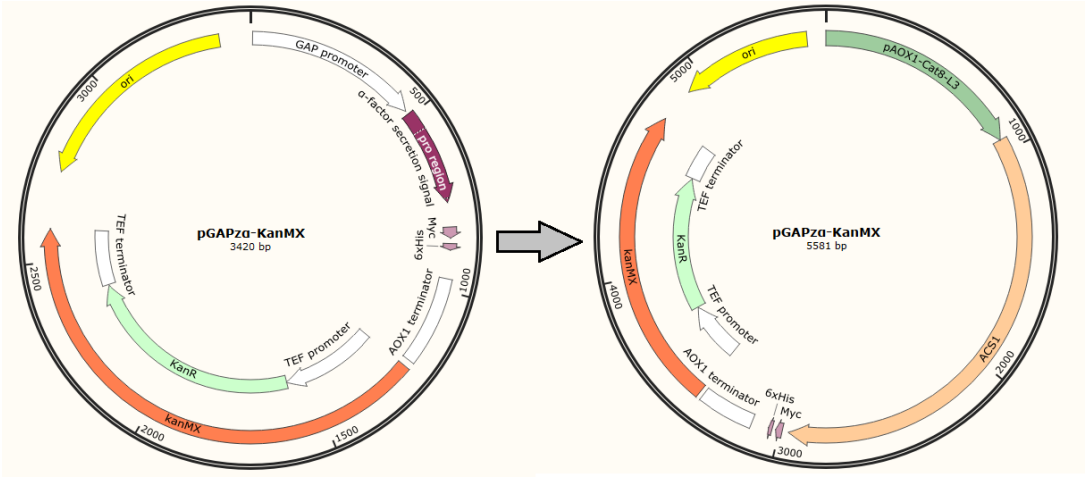


Figure 3. 4 Map of the constructed plasmid pAOX1-Cat8-L3:: *ACS1* from pGAPzα-A.

3.4.3 Construction of Recombinant Plasmids pmAOX1::*mApple*

The construction procedure of the plasmid pmAOX1::mApple started with primer design, in line with the detailed explanations in section 3.5.1.1. Related primers, which were designed for use in the amplification processes of desired genes, were supplied from Sentegen (Ankara, Turkey).

3.4.3.1 Amplification of the desired DNA

Polymerase chain reaction (PCR) was carried out to amplify the desired DNA fragments. Optimization of annealing temperatures and extension times to increase PCR efficiency were made based on designed primers, and PCR reactions were taken place in a thermocycler (Techne®, Flexigene, and TC-3000X). Amplifications of the *modified alcohol oxidase 1* promoter and mApple gene were provided from the pmAOX1::*eGFP* and pADH2-Cat8-L2::*mApple* plasmids, respectively. PCR operation conditions for amplification of PmAOX and mApple were given in Tables

3.23 and 3.24. Furthermore, the composition of PCR reaction mixtures for P_{mAOX1}-mApple was demonstrated in Table 3.25

Table 3. 23 *PCR operation conditions for amplification of pmAOX1*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	61.7 °C	30 sec	
Extension	72 °C	35 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 24 *PCR operation conditions for amplification of mApple*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	68.8 °C	35 sec	
Extension	72 °C	80 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 25 *Composition of PCR reaction mixtures for pmAOX1 and mApple*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template DNA	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

Afterward, of amplification of pmAOX1 and mApple fragments, PCR purification was carried out by using GeneJET PCR Purification Kit (ThermoFisher, USA) to eliminate the unwanted PCR elements and byproducts. Then, the fusion of the P_{mAOX1} fragment to mApple protein was achieved with the help of fusion PCR. Algorithm of construction logic of P_{mAOX1}::mApple fragment demonstrated in Figure 3.5.

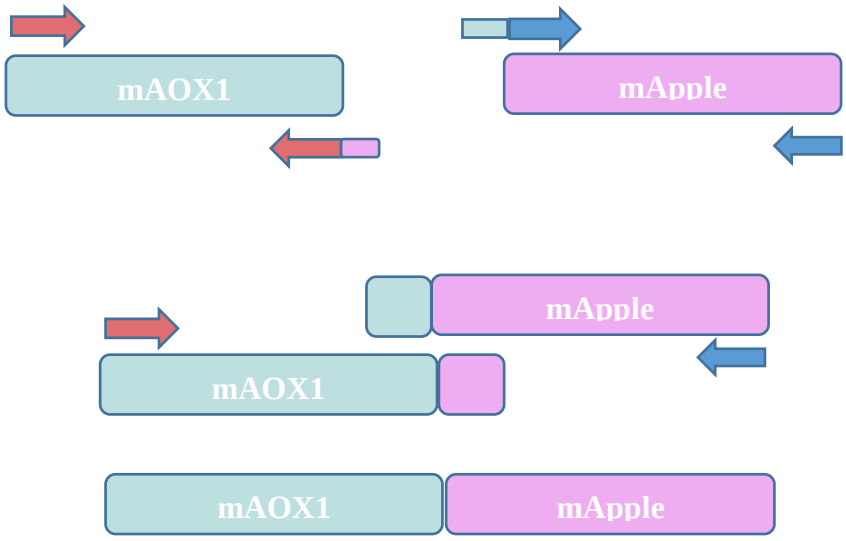


Figure 3. 5 Algorithm of construction logic of P_{mAOX1}::mApple fragment

By the usage of forward and reverse primers of pmAOX1 and mApple, amplification of both pmAOX1 and mApple fragments was accomplished, respectively. Reverse primer of pmAOX1 and forward mApple designed in such a way that they have overlapping sequences with each other to be able to fuse these two fragments. In a separate PCR application, these two fragments have coalesced, and the construction of the PmAOX1::mApple fragment is completed. Thermo-cyclic PCR conditions and reaction compositions for overlap extension showed in Table 3.26 and Table 3.27.

Table 3. 26 *Thermo-cyclic PCR operation conditions for construction of pmAOX1-mApple*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	68.8 °C	30 sec	
Extension	72 °C	60 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

Table 3. 27 *Composition of PCR reaction mixture for construction of P_{mAOX}-mApple*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template 1 (mAOX1)	2 µL
Template 2 (mApple)	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

3.4.3.2 Purification of the PCR Products and Digestion Reactions

For the elimination of meddling elements of the PCR reaction, agarose gel electrophoresis was performed on SOE product, pmAOX1-mApple, and purified by gel elution with the application of GeneJet Gel Extraction Kit (ThermoFisher, USA) following the instructions explained in section 3.4.5. Digestion with the same restriction enzymes, which designated in the primer design procedure, of both vector DNA, pADH2-Cat8-L2::*hGH*, and insert DNA, PmAOX1::mApple is crucial to construct the desired plasmid. Thus, double digestion of both vector and insert DNA was carried out with *Asc*I and *Xba*I restriction enzymes. The digestion process was taken place in the water bath at 37°C overnight. Composition of the digestion reaction mixture with *Asc*I and *Xba*I for insert and vector demonstrated in Table 3.28 and 3.29, respectively.

Table 3. 28 *Composition of the digestion reaction mixture with AscI and XbaI for vector, pADH2-Cat8-L2::hGH*

Component	**20 µL Reaction**
10X CutSmart Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 100ng/µl)	8 µl
*Asc*I	1 µl
*Xba*I	1 µl
H_2O	Up to 20 µl

Table 3. 29 *Composition of the digestion reaction mixture with AscI and XbaI for insert, PmAOX1::mApple*

Component	**20 µL Reaction**
10X CutSmart Buffer	2 µl
DNA (0.5-1µg) (DNA concentration is 50ng/µl)	12 µl
*Asc*I	0.6 µl
*Xba*I	0.6 µl
H_2O	Up to 20 µl

Following the double digestion of vector and insert, both vector and insert made run on agarose gel and extracted from the gel. Gel elution was carried out by the usage of the GeneJet Gel Extraction Kit (ThermoFisher, USA); as a result of this, purification of fragments was achieved.

3.4.3.3 Ligation reaction of the vector and the insert

T4 DNA ligase enzyme used for the ligation process of the vector and insert. For the ligation reaction, 20 µl of the reaction mixture was determined as the proper amount. The ligation process was carried out at 16°C for 16 hours, and the composition of the ligation mixture is tabulated in Table 3.30. The optimum ratio of the vector to insert DNA and quantity of the vector were determined as 3::1 and 50 ng, respectively. Furthermore, the quantity of the insert DNA was obtained from Equation 3.1.

Table 3. 30 *Composition of the ligation reaction mixture of vector and insert*

Component	**20 µL Reaction**
T4 DNA Ligation Buffer	2 µl
Vector DNA, pADH2-Cat8-L2::*hGH* (DNA concentration is 20 ng/µl)	2.5 µl
Insert DNA, PmAOX1:: mApple (DNA concentration is 10 ng/µl)	10.3 µl
T4 DNA Ligase	1 µl
H_2O	Up to 20 µl

Following ligation, the transformation process for *E. coli* transformation, which was explained in the previous part of this section, was carried out. With the ligation of the insert and vector and transformation to *E. coli*, the construction of the desired plasmid was accomplished and shown in Figure 3.4.

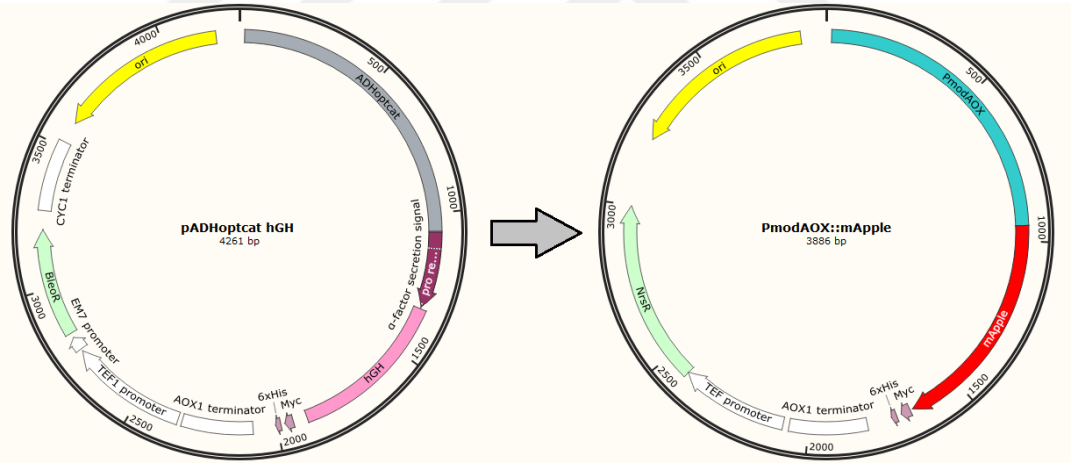


Figure 3. 6 Map of the constructed plasmid pmAOX1:: mApple from pADH2-Cat8-L2::*hGH*

Schematic representation of the metabolic engineering strategy for the construction of the desired plasmid is demonstrated in Figure 3.6.

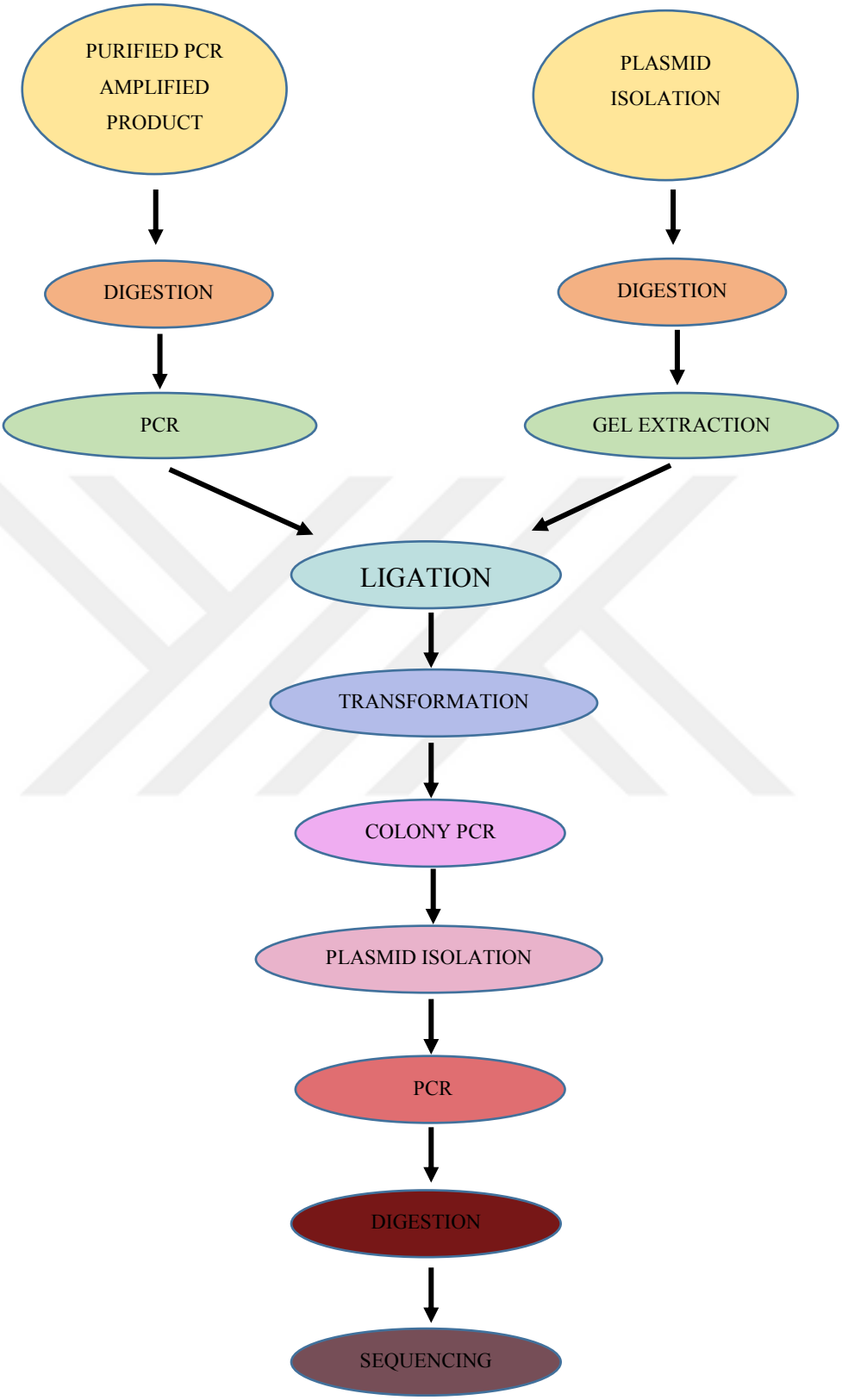


Figure 3. 7 Schematic representation of the metabolic engineering strategy for construction of desired plasmid

3.5 Determination of the gene-copy number

To compare the effects of overexpressed genes on cell metabolism and production capacity and to investigate their relationship with gene copy numbers, it is critical to know the copy amount of overexpressed genes in the strains constructed. In this context, the copy numbers of *ADH2*, *ACS1*, and mApple genes in strains created by metabolic engineering methods were investigated by relative qualification method using quantitative chain reaction. For this purpose, the argininosuccinate lyase, *ARG4* gene, which is known to be in a single copy in the *P. pastoris* genome, was selected, and the qualification of the desired genes was made comparatively on the basis of this gene. Firstly, to achieve this purpose, genomic DNA of constructed *P. pastoris* strains was isolated by using Wizard Genomic DNA Purification Kit (Promega)in the direction of instructions. Then, specific primers for qPCR from the inner part of the genomic sequences of the investigated genes and standard primers for amplification of the whole sequences of desired genes were designed with the help of Oligo Analyzer 3.1 and listed in table 3.31.

Table 3. 31 *Primers designed for the qPCR experiments and amplification of standards*

Name	Sequence (5'-3')	Length (bp)	Tm (°C)	GC Content	Size of Amplicon
ARG4-Std-F	GTTT ACA CTG AGG GCC TGG A	20	56.8	55	1259
ARG4-Std-R	GACTCTAGCTTTTCA TTCAGTGC	23	53.8	43.5	1259
ARG4-qPCR-F	GGTGAGTTGATTGGT CGTGG	20	62.7	55	185
ARG4-qPCR-R	CCGGGCATCAAGACG TCTAT	20	63.6	55	185

Table 3.31 (Continued)

ADH2-Std-F	CTTATTTACTTTACGAAAATG TCTCCAACTATCCCAACTACAC	43	60.6	35	1053
ADH2-Std-R	GATGGTACCTTATTTGGAAGTGTCCACAACG	31	60.5	45	1053
ADH2-qPCR-F	CCCCACATGCTATCAACCAATC	22	63.2	50	192
ADH2-qPCR-R	AACCTCTGGTGAACAAGTCGAT	22	63.8	45.5	192
ACS1-Std-F	CAACTAATTATTCGAAACGATGCCATTAGATAACGAACACTTACTTC	47	61.7	34	2007
ACS1-Std-R	GTTGGTACCTCATTTGCGGGCATC	24	60.6	54	2007
ACS1-qPCR-F	GATTGTCTACTGCCGAGATTG	21	60.4	47.6	206
ACS1-qPCR-R	TTTAGGTGCAGCGAATGG	18	60	50	206

Table 3.31 (Continued)

mApple-Std-F	**GAGAAGATCAAAAAACAACTAATTATTCGAAACGatggtgagcaagggcg**	**50**	**65.4**	**40**	**711**
mApple-Std-R	GCCTCTAGATTACTTGTACAGCTCGTCCATGCC	33	63.7	52	711
mApple-qPCR-F	GTCAAGACCACCTACAA	17	63	47	105
mApple-qPCR-R	CACGATGGTGTAGTC	15	61	53	105

For copy number determination, *ARG4* was used as a reference gene since just a single copy of it exists in the *P. pastoris* genome. Copy numbers of other genes, *ADH2*, *ACS1*, and *mApple,* in the selected strains were normalized relative to the *ARG4* gene. For this reason, standard curves of these four genes were essential to making an accurate relative quantification of copy numbers of desired genes. In this context, *ARG4*, *ACS1*, and *ADH2* gene sequences were amplified by conducting a PCR experiment in which isolated genomic DNA of wild-type *P. pastoris* was used as a template. Furthermore, the mApple gene sequence was amplified from the isolated pAHD2-Cat8-L2::*mApple* plasmid. For amplification of standards, primers that are designed from the external sites of the gene sequences were used so that amplifying the whole gene sequences of standard genes. Reaction compositions and thermocycling conditions of amplification of standard genes via PCR were given in Tables 3.32 and 33, respectively.

Table 3. 32 *Composition of PCR reaction mixture for standards*

Component	**50 µL Reaction**
5X Q5 Reaction Buffer	10 µL
5mM dNTPs	2 µL
10 µM Forward Primer	2.5 µL
10 µM Reverse Primer	2.5 µL
Template DNA	2 µL
Q5 High-Fidelity DNA Polymerase	0.5 µL
H_2O	Up to 50 µL

Table 3. 33 *PCR operation conditions for amplification of standards*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	98 °C	30 sec	1
Denaturation	98 °C	10 sec	30
Annealing	65.7 °C	30 sec	
Extension	72 °C	80 sec	
Final Extension	72 °C	2 min	1
Hold	4 °C	∞	-

After amplifications were completed, PCR products were purified with GeneJet PCR Purification Kit (Thermofisher). Afterward, to obtain the standard curves of the *ARG4*, *ADH2*, *ACS1*, and *mApple* genes, dilutions of 103, 104, 105, 106 copy numbers were prepared. For this purpose, firstly, by the application of the following formula, copy quantities of the standard samples were calculated.

$$copy\ number = \frac{6.022 \times 10^{23} \times amount\ of\ sample}{Molecular\ weight\ of\ DNA \times 10^{9}}$$

Amplicon molecular weights of the standards were obtained from the Bioinformatics online DNA molecular weight calculator (http://www.bioinformatics.otg/sms2/dna_mw.html) as follows:

- Amplicon Molecular Weight of *ARG4* Standard: 777867.71
- Amplicon Molecular Weight of *ADH2* Standard: 650681.25
- Amplicon Molecular Weight of *ACS1* Standard: 1240045.15
- Amplicon Molecular Weight of *mApple* Standard: 439436.87

Then, genomic DNA concentrations were diluted to a concentration of 2 ng / µl, and 9 µl of diluted genomic DNA was used as a template for each qPCR reaction. Moreover, qPCR primers that were designed from the inner side of the gene sequences were also used in the qPCR experiment. The qPCR was performed on the LightCycler® 480 Instrument II (Roche) instrument using the LightCycler® 480 SYBR Green I Master (Roche) mix. qPCR reaction mixture compositions and thermo-cycler operation conditions were tabulated in Tables 3.34 and 35, respectively.

Table 3. 34 *Composition of qPCR reaction mixture*

Component	**20 µL Reaction**
Master Mix 2X Concentration	10 µL
10 µM Forward Primer	0.5 µL
10 µM Reverse Primer	0.5 µL
Template DNA	9 µL

Table 3. 35 *Thermocycler operation-profile of qPCR experiments*

Steps			Time	Cycle
Denaturation	Initial Denaturation	95 °C	10 min	1
Amplification	Denaturation	95 °C	10 sec	45
	Annealing	58 °C	5 sec	
	Elongation	72 °C	10 sec	
Melting	Melting	50 °C to 99°C	Continuos with slope of 1°C/s	1
Cooling	**Keep**	40 °C	30 sec	1

3.6 Growth Media

Sterilization of the growth media in this study was achieved either by autoclave at 121°C for 20 minutes or filtered (22µM).

3.6.1 Solid Medium

The growth of *E. coli* cells from glycerol stock was carried out on the LB agar plate with the appropriate antibiotic at 37°C for 24 hours. 150 µg/ml Hygromycin was selected for the *E. coli* strains, which carries the for P_{ADH2}:: *ADH2* and 50 µg/ml nourseothricin (NTC) was selected for the strains, which carries the P_{mAOX1}::mApple, for antibiotic resistance. For the wild-type *E. coli* DH-5α cells, any antibiotic was not added to the agar. Furthermore, for the growth of *P. pastoris* cells, they were inoculated onto the YPD agar containing proper antibiotics and incubated at 30°C for 48 hours. 200 µg/ml Hygromycin was used for the *P. pastoris* strains, which carries the for P_{ADH2}:: *ADH2* and 50 µg/ml nourseothricin (NTC) was used for the strains, which carries the P_{mAOX1}::mApple. For the wild-type *P. pastoris* X-33

cells, there was no need for an antibiotic, and YPD agars were prepared without antibiotics. The composition of the LB and YPD agar medium was tabulated in Tables 3.36 and 3.37.

Table 3. 36 *Composition of LB Agar Medium for E. coli*

Compound	**Concentration (g/L)**
Tryptone	10
Yeast extract	5
NaCl	5
Agar	15

Table 3. 37 *Composition of YPD Agar Medium for P. pastoris*

Compound	**Concentration (g/L)**
Yeast extract	10
Peptone	20
Dextrose	20
Agar	20

3.6.2 Precultivation Medium

Inoculation of *P. pastoris* cells, which grow on the YPD agar, into the precultivation medium, BMGY or YP, is necessary. Precultivation process can be carried out either in 12 deep-well plates or air-filtered shake bioreactors. The composition of the BMGY and YP media were listed in Table 3.38 and Table 3.39.

Table 3. 38 *Composition of the BMGY for the precultivation of P. pastoris cells*

Compound	**Concentration (g/L)**
Yeast extract	10
Peptone	20
Glycerol	10 ml
Yeast Nitrogen Base (YNB, w/o amino acids)	4.08
Potassium phosphate buffer (pH=6)	0.1 M
Ammonium sulfate	12
Biotin	0.0004
Chloramphenicol (34 mg/ml)	1 ml

Table 3. 39 *Composition of the YP for the precultivation of P. pastoris cells*

Compound	**Concentration (g/L)**
Yeast extract	10
Peptone	20

3.6.3 Fermentation Media

After completion of the precultivation, cells were harvested and inoculated into fermentation media, which have different carbon sources, in a way that the initial cell concentration was equal to 1 OD600. Two different media were selected for the screening experiments, and their compositions were given in Tables 3.40 and 3.41.

Table 3. 40 *Defined fermentation medium for screening*

Compound		**Concentration (g/L)**
$(NH_4)_2HPO_4$		4.95
$MgSO_4.7H_2O$		14.9
$CaCl_2.2H_2O$		1.17
PTM Trace Element Solution		1.47 mL
Biotin (0.2 g/L)		2 mL
Potassium phosphate buffer (pH=6)		0.1 M
Carbon source	**Name of the Condition**	**Concentration**
	Excess Glucose	20 g/L
	Excess Glycerol	20 g/L
	Limited Glucose	2 g/L
	Ethanol	1 % (v/v) or 2 % (v/v)
	Methanol	2 % (v/v)

Table 3. 41 *ASMV6 Defined cultivation-base-fermentation medium for screening*

Compound	**Concentration (g/L)**
$(NH_4)_2HPO_4$	6.3
$(NH_4)_2SO_4$	0.8
$MgSO_4.7H_2O$	0.49
KCl	2.64
$CaCl_2.2H_2O$	0.0535
Citric acid monohydrate	22
PTM Trace Element Solution	1.47 mL
NH_4OH (25%)	20 mL
Biotin (0.2 g/L)	2 mL

Table 3.41 (Continued)

KOH solid	**Adjustment of pH 6.2-6.6**	
	Name of the Condition	**Concentration**
Carbon source	Excess Glucose	20 g/L
	Excess Glycerol	20 g/L
	Limited Glucose	2 g/L
	Ethanol	1 % (v/v) or 2 % (v/v)
	Methanol	2 % (v/v)

3.7 Screening Conditions

3.7.1 mApple and eGFP synthesis from selected r-*P. pastoris* strains

Intracellular synthesis of mApple and eGFP and the comparison of designed double promoter systems, as well as single and multi-copies of them, were conducted in 12-well deep plates. Selected colonies were inoculated into 2 ml of YP medium containing 25 µg / µL of zeocin or 50 µg / µL NTC, and incubated at 30°C for 18-20 hours at 200 rpm. Then, the cells were transferred to the defined fermentation medium and ASMv6 production medium, which were tabulated in Table 3.24 and 3.25, respectively, which contain selected carbon sources. Screening experiments of the production of the fluorescence proteins, mApple and eGFP, of constructed double promoter expression systems and their single copies, was carried out in a production medium, whose OD600 was equal to 1with 20 g / L glucose or glycerol, 2 g/L limited glucose, 2% (v / v) methanol and 2% ethanol (v / v). When fermentation time reached 22-24 hours, cells were harvested with centrifugation at 4500g for 5 minutes and diluted or concentrated to OD600 of 8. Afterward, measurement of mApple and eGFP production by the usage of these prepared solutions was

conducted with a fluorescence spectrophotometer (Agilent technologies). Detection of eGFP and mApple experiments was carried out in duplicate.

3.8 Analyses

3.8.1 Cell concentration

Wet cell concentrations (g/L) of the samples, which were collected during the exponential phase, was calculated by applying the following formula:

$$C_x = 0.24 \times OD_{600} \times Dilution\ factor \qquad (3.2)$$

3.8.2 mApple and eGFP synthesis

Measurements of the production of mApple and eGFP were carried out with Fluorescence Spectrophotometer (Agilent technologies) by the usage of 96-well microtiter-plates (Thermoscientific). Measurements of the mApple (red fluorescence protein) were accomplished at the excitation and emission wavelengths of 568 nm and 592 nm, respectively. On the other hand, measurements of eGFP (green fluorescence protein) were carried out at the excitation wavelength of 488 nm and emission wavelength of 509 nm. The cell suspension was either diluted to OD_{600}=8 or concentrated to OD_{600}=8 to conduct the measurements and mixed with pipetting by the usage of a multi-pipette. For normalization of the eGFP and mApple production measurements, the background signal of the wild-type *P. pastoris* X-33 cells was removed. According to the intensity value, which was read from the fluorescence spectrophotometer, the production level of mApple and eGFP were given in arbitrary units (a.u.).

3.8.3 Ethanol concentration measurements

Extracellular ethanol concentrations in the medium were determined by gas chromatography (Agilent GC 6850, Wilmington, USA) equipped with a 30 m x 530 µm x 40 µm Agilent 19095P-Q04E column and operated at a constant oven temperature of 170 °C. Helium gas, given at a constant pressure of 19.62 psi and a flow rate of 120 mL/min, was used as the carrier gas. The TCD detector was operated at 200 °C and 5 Hz with a reference flow of 12 mL/min and analyzed for 3 minutes by injecting 1 µL of the sample.

3.8.4 Extracellular metabolite concentration

High-pressure liquid chromatography (Waters, Alliance 2695, Organic Acid Analysis System) was used to determine the concentrations of organic acids, particularly acetaldehyde and acetic acid, secreted by the microorganism during the air-filtered shake-flask bioreactor experiments. Organic acid concentrations were calculated with the help of standard chromatograms obtained by standard organic acid analyzes of the known concentrations. Calibration curves of related metabolites were given in Appendix. Before loading samples to the HPLC system, they were filtered with 45µm filters. The analysis conditions of organic acids are as follows (Ileri, 2006):

Column	: Capital Optimal ODS, 5 µm
Column dimensions	: 4.6 mm x 250 mm
System	: Reverse phase chromatography
Mobile phase flow rate	: 0.8 ml/min
Column temperature	: 30°C

Detector, wavelength ve sensitivity : Waters 2487 Dual absorbance detector, 254 nm

Injection volume : 5 µl

Analysis duration : 15 dk

Mobile phase content : 0.312% (w/v)) NaH_2PO_4 ,0.062% (v/v) H_3PO_4

3.9 Total RNA isolation

To be able to conduct transcript level analysis with constructed *P. pastoris* strains, total RNA isolation was performed for each strain. For this purpose, firstly, samples were collected 2 hours after starting the exponential growth phase. The amount of sample taken was designated as 4×10^8 cells/1ml, which was approximately equal to 8 OD. (Asada *et al*., 2011; Yuan *et al*., 2021). Then, taken cells were harvested by centrifuging at 4500×g for 5 minutes at 4°C. Afterward, QIAGEN RNaesy Mini Kit was used for total RNA isolation in accordance with instructions as follows: firstly, harvested cells were resuspended with 600 µl buffer RLT and, the mixture was gently vortexed. Afterward, 500µl acid-washed glass beads were added into the mixture, and to disrupt the cell membrane, tissue-lyser was used at the maximum speed of 30Hz for 1 minute. After 1 minute was completed, 20 seconds waited to prevent the machine from overheating. This step was repeated five times. Then, two minutes waited at room temperature for letting beads going down. Finally, approximately 350µl of the sample was taken into a new RNAse-free microcentrifuge tube and centrifuged for 2 min at full speed. Then the supernatant was transferred to a new RNase free microcentrifuge tube, so sample preparation was completed. RNA isolation part started with the addition of 1 volume (approximately 350µl) 70% ethanol into supernatant and mixing it by pipetting. Then, the sample (approximately 700 µl) was transferred to an RNaeasy spin-column

collection tube, which was supplied by the kit, and centrifuged at 8000-1000g for 15 seconds. Afterward, flow-through was discarded, and 350 µl buffer RW1 was added into the RNaeasy spin column. Another centrifugation at 8000-10000g for 15s was applied to wash the spin column membrane, and flow-through was discarded. Thuswise, precipitation of nucleic acids was completed. Thereafter, to remove DNA from the supernatant, 3 µl DNaseI 30 µl buffer and 67 µl RNase-free water were mixed in a fresh RNase-free microcentrifuge tube and homogenized by gentle inversion of the tube. Then, this DNaseI incubation mixture (100 µl) was directly added onto the RNaesy spin column membrane. The sample was kept at room temperature for 15 minutes, and DNA removal was completed. To wash RNA, 350 µl buffer RW1 was added to the RNaesy spin column and centrifuged for 15 seconds at 8000-10000g. Flow-through was discarded, and 500 µl RPE buffer was added onto the spin column, and another centrifugation was performed at 8000-10000g for 15 seconds. Flow-through was discarded, and again 500 µl RPE buffer was added onto the spin column. This time, the sample was centrifuged for 2 minutes at 8000-10000g, and the spin column was placed in a fresh RNase-free 2 ml collection tube (supplied). The spin column was centrifuged for 1 minute at 8000-10000g. Next, for the elution of RNA, the RNaeasy spin column was placed in a 1.5 ml fresh RNase-free microcentrifuge collection tube (supplied), and 35 µl RNase-free water was added directly onto the spin column membrane. The tube was incubated for 2 minutes at room temperature and centrifuged for 1 minute at 8000-10000g for the elution of RNA. Finally, elute was taken and added again onto the spin-column membrane. 2 minutes of incubation was performed. The sample was centrifuged at 8000-10000g for 1 minute. Thus, the procedure was completed, and high concentrated RNA was obtained. The concentration of isolated RNA was measured by Nanodrop.

3.9.1 Transcript level analysis

The threshold cycle (ΔΔ*CT*) approach was used to calculate changes in transcript levels in *P. pastoris* mutant strains compared to the wild-type control. The expression of the actin gene ACT1 was used to standardize all results. Primers used in the qRT-PCR analysis were given in Table 3.42.

Table 3. 42 *List of primers used for the transcript level analysis.*

Primer	Sequence (5'-3')	Used for
Act1_qRT PCR Fw	TGAGGCTTTGTTCCACCCAT	*Actin* mRNA quantification
Act1_qRT PCR Rv	CTCAGCAATACCTGGGAACA TAGTAG	*Actin* mRNA quantification
ADH2_qRT-PCR_Fw	CCCCACATGCTATCAACCAAT C	*ADH2* mRNA quantification
ADH2_qRT-PCR_Rv	GATCTCGATGGACTTCAGAAC GT	*ADH2* mRNA quantification
ALD4_qRT-PCR_Fw	CGGTCTTGCTGCTGGTAT	*ALD4* mRNA quantification
ALD4_qRT-PCR_Rv	TTTGGTGGAAATCGTTGTAGG T	*ALD4* mRNA quantification
ACS1_qRT-PCR_Fw	GGTTTTGGCTGGAGAGGAAG A	*ACS1* mRNA quantification
ACS1_qRT-PCR_Rv	TTGCGGGCATCCCTTTT	*ACS1* mRNA quantification

Table 3.42 (Continued)

FBP1_qRT-PCR_Fw	TGAAGAAACCCCAAGCAAAC	FBP1 mRNA quantification
FBP1_qRT-PCR_Rv	TGGAGTCTGCTGGATAGC	FBP1 mRNA quantification
ICL1_qRT-PCR_Fw	CAGAAATGGTCAGGAGCCG	ICL1 mRNA quantification
ICL1_qRT-PCR_Rv	AATGGTCCTTGAATTGATCTTCAG	ICL1 mRNA quantification
MLS1_qRT-PCR_Fw	ACTGGCGAGAAGATCACA	MLS1 mRNA quantification
MLS1_qRT-PCR_Rv	CTCCACTAATCTCCTTCTTCAAAC	MLS1 mRNA quantification
DAS1_qRT-PCR_Fw	CGGTAAGTCTCTTCCTGTTG	DAS1 mRNA quantification
DAS1_qRT-PCR_Rv	TTGGTTTTCCCTTCAAGTCG	DAS1 mRNA quantification
Mxr1_qRT-PCR_Fw	ACAATGACACAAATGCTGCTGATGCT	Mxr-1 mRNA quantification
Mxr1_qRT-PCR_Rv	GCATGAACAGGCGGTCTGAATCG	Mxr-1 mRNA quantification
CAT8-1_qRT-PCR_Fw	TAATGTATCTCCTCCCAATAGTGAAAG	Cat8-1 mRNA quantification
CAT8-1_qRT-PCR_Rv	GTCCGAATAAACTCCCAGCAG	Cat8-1 mRNA quantification

Table 3.42 (Continued)

CAT8-2_qRT-PCR_Fw	GGTCTTCAGCGTCTGGTT	Cat8-2 mRNA quantification
CAT8-2_qRT-PCR_Rv	ACTTTTCTGAGCTTTTCACCTG	Cat8-2 mRNA quantification
PDA1_qRT-PCR_Fw	TGGAATATCTACAGAGGAGGAAC	PDA1 mRNA quantification
PDA1_qRT-PCR_Rv	TCAGCTTCTTTGGTTTGCTT	PDA1 mRNA quantification
PGK1_qRT-PCR_Fw	TAAGAAGTTCGGTGGTGCTGACA	PGK1 mRNA quantification
PGK1_qRT-PCR_Rv	CCCTCCAACAGTTCCAAAGAAGC	PGK1 mRNA quantification
CDC19_qRT PCR Frw	GCCCAATCCTTATGGTAACC	CDC19 mRNA quantification
CDC19_qRT PCR Rv	GGCTTCAGTGATGGCATATTG	CDC19 mRNA quantification
PYC2_qRT-PCR_Fw	GGCGAAGTCAAGAAGGGTGAT	PYC2 mRNA quantification
PYC2_qRT-PCR_Rv	TCGTTCTCCTTGACTGCGAT	PYC2 mRNA quantification

CHAPTER 4

RESULTS AND DISCUSSION

The objective of this study is to increase the ethanol uptake rate and decrease by-product formation in the ethanol utilization (EUT) pathway and acquire a smoothly operating intracellular reaction network from the TCA cycle to the gluconeogenesis pathway through anaplerotic reactions. For this purpose, a novel metabolically engineered host, *Pichia pastoris*, was developed by overexpression of the gene(s) of the enzyme(s) that catalyzes the limiting reactions in the EUT pathway under a hybrid-structured synthetic promoter. The effects of the engineering in the EUT pathway in *P. pastoris* metabolism were investigated. Ensuring that the smooth operation of EUT pathway and the intracellular network of the microorganism; to increase therapeutic protein production from ethanol by green-de novo-processes is another aim of this study.

In this context, in the first part of the work, the alcohol dehydrogenase 2 (ADH2), and Acetyl-CoA synthetase (ACS1), two enzymes that potentially catalyze the rate-limiting reactions, was overexpressed single handedly and together to find out how they affect the ethanol uptake rate, cell growth, and byproduct formation. Subsequently, the rate-limiting step in the EUT pathway, r-protein production under the engineered *P. pastoris* strains were investigated. In the second part of the work, effects of *ADH2* and *ACS1* overexpressions, one by one and together, on transcription levels in central metabolism were investigated. In the third part of the work, the effects of *ACS1* overexpression on r-protein production was investigated.

4.1 Effect of ethanol on EUT pathway

The effect of ethanol on the EUT pathway was investigated in air-filtered shake bioreactors with wild-type *P. pastoris* X-33. *The* strain was inoculated with an initial OD600 = 1.0 (Cx = 0.24 g/L) into the ASMv6 medium prepared (Table 3.44) with the carbon-source 2% (v/v) ethanol. At t = 36 h, the concentrations of the excreted acetaldehyde and acetic acid, which are the by-products of the EUT pathway, were presented in Figure 4.1.

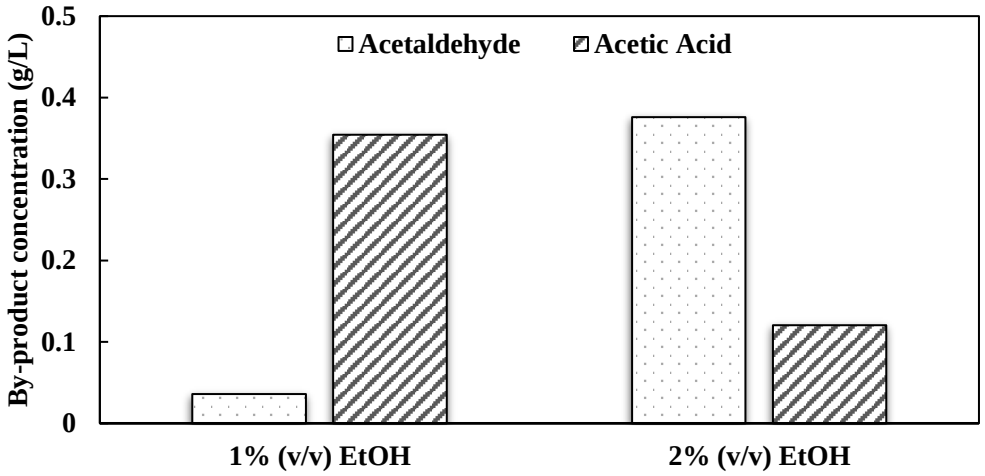


Figure 4. 1 By-product concentrations in the cell culture at t = 36 h excreted from wild-type *P. pastoris* X-33 strain on 1% (v/v) and 2% (v/v) ethanol

At 1% (v/v) initial ethanol concentration, at t = 36 h of the bioprocess extracellular C_{AcAl} and C_{AcAc} were determined as 0.19 and 0.35 g/L. When ethanol was used as the carbon source at an initial concentration of 2% (v/v), due to the pressure caused by high ethanol concentration C_{AcAl} and C_{AcAc} were 0.37 and 0.12 g/L, respectively. At initial ethanol concentration 1% (v/v), the by-product acetic acid concentration in the fermentation broth was 2.53-fold higher than the acetaldehyde concentration. When initial ethanol concentration was increased to 2% (v/v), the acetaldehyde concentration was 3.25-fold higher than the acetic acid concentration. The stress

caused by higher ethanol concentration altered the rate-limiting reaction of EUT pathway and changed the concentration ratios of the by-products. Based on the by-product accumulation data, it is concluded that one of the three consecutive reactions in the EUT pathway should be the potential the rate-limiting step. Therefore, first and third genes on the EUT pathway (Figure 4.2) were overexpressed single-handedly and then together to determine the effects of the encoded ADH2 and ACS1 enzyme concentrations on the EUT pathway.

4.2 Overexpression of the enzymes in the EUT pathway

4.2.1 Overexpression of Alcohol dehydrogenase 2 (ADH2) enzyme

In the EUT pathway, ethanol is converted to acetyl coenzyme-A (Acetyl-CoA) over three consecutive reactions (Figure 4.2). Then, acetyl CoA transfers its acetyl group to oxaloacetate to form citrate. The ADH2 enzyme is the first enzyme in the EUT pathway that catalyzes the oxidation of ethanol to acetaldehyde.

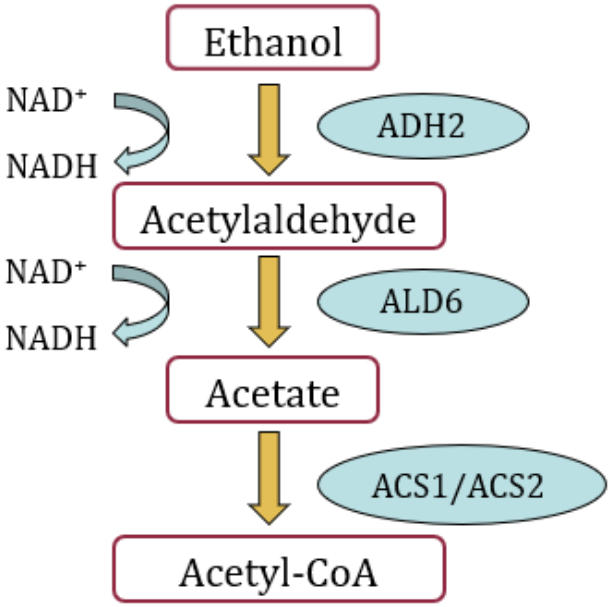


Figure 4. 2 Ethanol utilization pathway in *Pichia pastoris*

4.2.1.1 Construction of the recombinant plasmid and strains carrying pADH2-wt::*ADH2*

Aiming the overexpression of ADH2 enzyme in the EUT pathway, a recombinant plasmid containing pADH2-wt::*ADH2* cassette was constructed. For this purpose, pADH2-wt::*ADH2* cassette was inserted into the vector-plasmid pGAPZα-A. For the construction of plasmid, proper restriction enzymes, which are non-cutter of insert and one-site cutter of vector plasmid, *Kpn*I and *Bgl*II were chosen. SnapGene and Restriction mapper was used for finding the required restriction sites. In Figure 4.3, vector plasmid pGAPZα-A, and the restriction enzymes used in the construction of recombinant plasmid are presented.

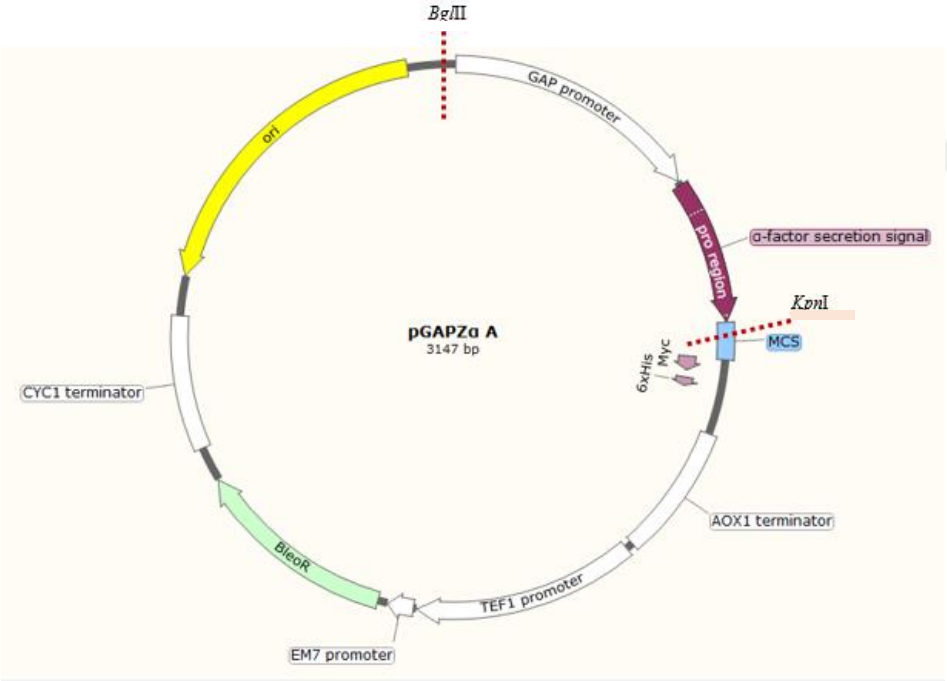


Figure 4. 3 Base plasmid pGAPZα-A, and restriction enzymes *Bgl*II and *Kpn*I, used in construction of recombinant plasmid.

In the parent pGAPZα-A plasmid, there is a *Bgl*II recognition site at the 5' end of pGAP gene (Figure 4.3); however, *Bgl*II cuts the pADH2-wt gene nucleotide

position 744 bp, so insert cannot be digested with *Bgl*II. Besides *Bgl*II, another RE is required for the digestion of the insert pADH2-wt::*ADH2* (Figure 4.4). Ligation is possible between both identical ends and between the different but compatible ends. New restriction sites can be generated by ligation of DNA fragments with compatible cohesive ends by cleavage with two restriction endonucleases that produce compatible overhangs. *Bam*HI and *Bgl*II generate different but compatible ends, which makes the ligation possible, and after ligation, the newly formed sequence can be recognized by neither *Bam*HI nor *Bgl*II. Thus, the *Bam*HI RE recognition region has been added to the pADH2-wt 5' end to be able to ligate the vector pGAPZα-A, and insert pADH2::*ADH2*. (Figure 4.4) Since *Bam*HI cannot cut the newly formed restriction site, it is still a one-site cutter of plasmid and can be used for linearization of plasmid before transformation to *P. pastoris* from AOX termination locus.

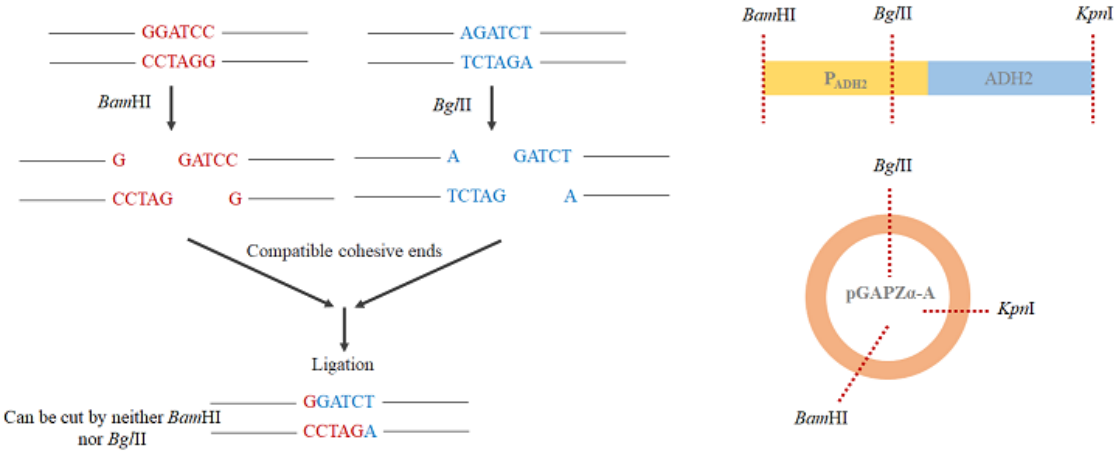


Figure 4. 4 Schematic representation of restriction endonucleases used in the construction of pADH2-wt::*ADH2* plasmid

pADH2-wt::*ADH2* cassette was constructed with Forward_$P_{ADH2\text{-}wt}$ and Reverse_ADH2 primers by PCR application. pADH2(1047 bp) and *ADH2* (1053 bp) were amplified by Q5 High-Fidelity DNA Polymerase. Wild-type *P. pastoris* genome was used as the template for amplification of the insert, pADH2-wt::*ADH2*.To optimize PCR reaction conditions, the effects of different annealing

temperatures were investigated for pADH2::*ADH2* cassette amplification. PCR reaction compositions and operation conditions were given in Tables 3.3 and 3.4, respectively.

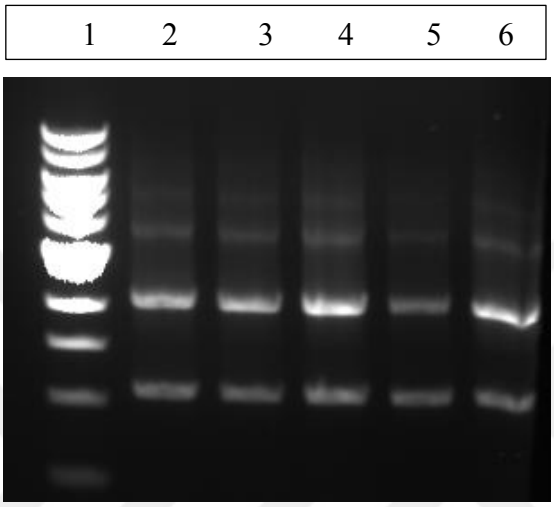


Figure 4. 5 Agarose gel electrophoresis image of the gene sequences amplified with Forward_$P_{ADH2\text{-}wt}$ and Reverse_*ADH2* primers at different annealing temperatures. 1: Gene Ruler Express (Thermoscientific); Annealing temperatures: 2: T= 60.8°C; 3: T = 62.7° C; 4: T = 65.7° C; 5: T = 67.8° C; 6: T = 69.2° C

The highest pADH2-wt::*ADH2* gene fragment concentration was obtained when the annealing temperature was T = 65.7°C (Figure 4.5). Thus, T = 65.7°C was determined as the optimum annealing temperature for amplification of pADH2-wt::*ADH2* fragment, and for further amplification, it was used as the optimum annealing temperature. The amplified fragment of pADH2-wt::*ADH2* was purified with GeneJet Gel Extraction Kit (ThermoFisher, USA) by separating from the other DNA fragments amplified. To obtain purified pADH2-wt::*ADH2* fragment, gel extraction was conducted. The gel electrophoresis image is presented in Figure 4.6.

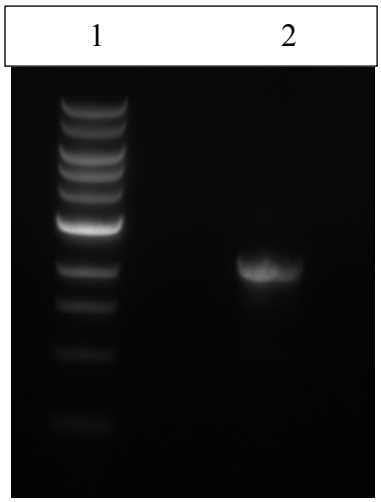


Figure 4. 6 Agarose gel electrophoresis image of the purified DNA fragment of pADH2-wt::*ADH2*. 1: Quick-Load Purple 1 kb DNA ladder; 2: Purified DNA fragment of pADH2-wt::*ADH2*

After the purification, the base-plasmid pGAPZαA, was digested sequentially with the restriction enzymes *Kpn*I and *Bgl*II, respectively, since *Kpn*I and *Bgl*II cannot work in a single buffer with high efficiency.

After the first restriction endonuclease digestion, pADH2-wt::*ADH2* fragment was purified with the PCR purification kit, and then a second restriction endonuclease digestion reaction was performed. *Bam*HI and *Kpn*I can work with high efficiency in a single and the same buffer, so inert pADH2-wt::*ADH2* was double digested with the restriction enzymes *Bam*HI and *Kpn*I in a single reaction mixture, and visualized in agarose gel electrophoresis after the digestion (Figure 4.7.). pADH2-wt::*ADH2* fragment was 2100 bp and fragments of double digested-plasmid pGAPZαA were 797 bp and 2350 bp (Figure 4.7).

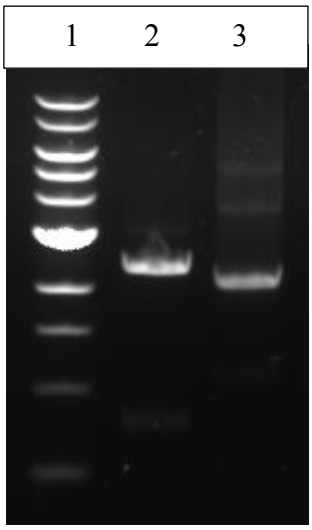


Figure 4. 7 Agarose gel electrophoresis of double digested pADH2-wt::*ADH2* fragment and plasmid pGAPZαA 1: Quick-Load Purple 1 kb DNA ladder; 2: Digested insert with *Bam*HI and *Kpn*I; 3: Digested base-plasmid, pGAPZαA, with *Kpn*I and *Bgl*II.

After the second digestion reaction of the base-plasmid and double digestion reaction of the insert, extraction of the DNA fragments from the gel was performed by razor procedure, purified with gel elution kit, and visualized in agarose gel electrophoresis after digestion (Figure 4.8). pADH2-wt::*ADH2* fragment is 2100 bp and the double digested vector plasmid pGAPZαA is 2350 bp (Figure 4.8).

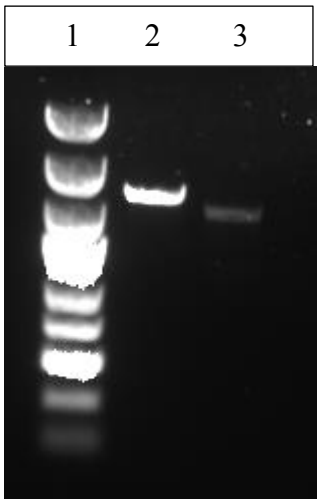


Figure 4. 8 Agarose gel electrophoresis of purified double digested pADH2-wt::*ADH2* fragment and plasmid pGAPZαA 1: GeneRuler Express; 2: Digested insert with *Bam*HI and *Kpn*I after gel elution; 3: Digested base-plasmid, pGAPZαA, with *Kpn*I and *Bgl*II REs after gel elution.

The ligation reaction was performed to clone the double digested pADH2-wt::*ADH2* into double digested vector pGAPZαA using the vector: insert 3:1 ratio under the conditions specified in Section 3.4.1.4.

After completion of the ligation reaction, by the application of the calcium chloride method, transformation of the constructed plasmid into wild-type *E. coli* DH5α cells was performed. To be able to make the selection, inoculation was done onto an agar medium that contains zeocin (Table 3.39). A single colony of *E. coli* DH5α transformed with pGAPZαA plasmid ligated with pADH2-wt::*ADH2* were used to inoculate 50 ml LB medium containing 25 μg/ml zeocin in 250 ml.

After 16-18 hours of inoculation at 37°C, six colonies were selected from putative transformants. Colony PCR, with Forward_P$_{ADH2\text{-}wt}$ and Reverse_ADH2 primers, was carried out to control if the integration of plasmid was achieved and visualized in agarose gel electrophoresis after the colony PCR (Figure 4.9). The pADH2-wt::*ADH2* fragment was 2100 bp. Taq DNA Polymerase (Thermo Scientific) was used for the control PCR. The PCR operation conditions, and reaction mixture contents of the colony PCR are presented in Table 4.1 and 4.2, respectively.

Table 4. 1 *PCR operation conditions for colony PCR*

Steps	**Temperature**	**Time**	**Cycle**
Initial Denaturation	95 °C	3 min	1
Denaturation	95 °C	30 sec	30
Annealing	62 °C	30 sec	
Extension	72 °C	2 min 20 sec	
Final Extension	72 °C	10 min	1
Hold	4 °C	∞	-

Table 4. 2 *Composition of PCR reaction mixture for colony PCR*

Component	**25 µL Reaction**
10X Taq Buffer	2.5 µL
5mM dNTPs	1 µL
10 µM Forward Primer	1.25 µL
10 µM Reverse Primer	1.25 µL
Template DNA	2 µL
Taq DNA Polymerase	2 µL
25mM $MgCl_2$	2 µL
H_2O	Up to 25 µL

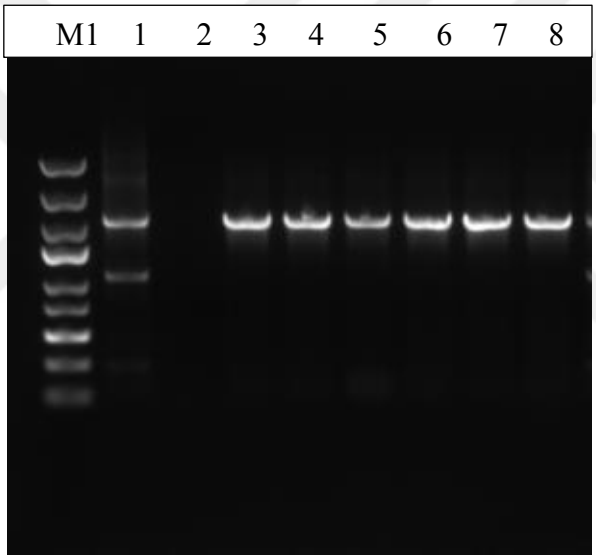


Figure 4. 9 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. M1: GeneRuler Express; 1: Positive control for PCR control of insert pADH2-*ADH2* gene; 2: Negative control for PCR control of insert pADH2-*ADH2* gene, 3-8: PCR products of potential plasmids

The results of the six selected colonies demonstrate that integration of the plasmid is correct (Figure 4.9). Plasmid isolation was performed for six positive putative transformants to confirm the achievement of insertion of constructed plasmid, and visualized in agarose gel electrophoresis after plasmid isolation (Figure 4.10). The size of the circular pADH2-wt::*ADH2* plasmid is 4467 bp.

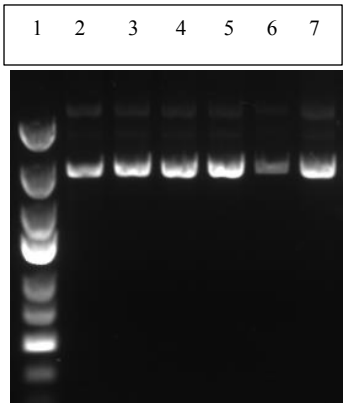


Figure 4. 10 Agarose gel electrophoresis of potential recombinant plasmids. 1: Generuler Express, 2-7: Potential plasmids carrying the $P_{ADH2\text{-}wt}$::*ADH2* fragment

Colony PCR, in which isolated plasmids were used as a template, was carried out with Forward_$P_{ADH2\text{-}wt}$ and Reverse_ADH2 primers under the condition given in Table 4.1 and 4.2, and visualized in agarose gel electrophoresis (Figure 4.11).

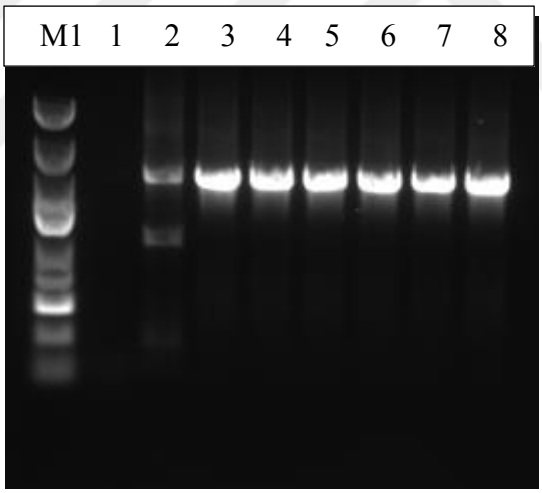


Figure 4. 11 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. M1: GeneRuler Express; 1: Positive control for PCR control of insert pADH2-*ADH2* gene; 2: Negative control for PCR control of insert pADH2-*ADH2* gene, 3-8: PCR products of potential plasmids

After colony PCR, two colonies were selected and the constructs between *Bgl*II and *Kpn*I restriction sites were verified with DNA sequencing analysis (Sentebio, Ankara) (Appendix B). Glycerol stocks of *E. coli* DH5α strains that carry the verified constructed plasmid, pADH2-wt::*ADH2*, were prepared and stored at -80°C.

The zeocin antibiotic resistance gene was replaced with the hygromycin encoding-gene in pADH2-wt::ADH2. As in the following stages of the study, combining different overexpression cassettes by sequential transformation is required to investigate the concerted effects of metabolic engineering in recombinant protein production. For this purpose, it was aimed to combine the constructed strains with DPESs with in the following stages. As a result, the necessity of using an antibiotic resistance gene different from Zeocin and NTC is required since DPESs carry Zeocin and NTC resistance genes.

The restriction enzymes, that could cut the zeocin gene sequence, including its own promoter (TEF) and terminator (CYC1), and could not cut the hygromycin sequence were identified as *Bam*HI and *Pci*I using SnapGene and Restriction mapper. Secondly, Hygromycin fragment (1479bp) with TEF promoter and TEF terminator were amplified with Forward_Hyg and Reverse_Hyg primers in accordance with the conditions given in Tables 3.9 and 3.10, respectively. Then, the amplified insert Hygromycin fragment was purified with PCR purification kit. Purified fragment was double digested with the restriction enzymes of *Bam*HI and *Pci*I in a single restriction reaction. Then, purified with PCR purification kit, and visualized in agarose gel electrophoresis. (Figure 4.12) Hygromycin with with TEF promoter and TEF terminator is 1479 bp.

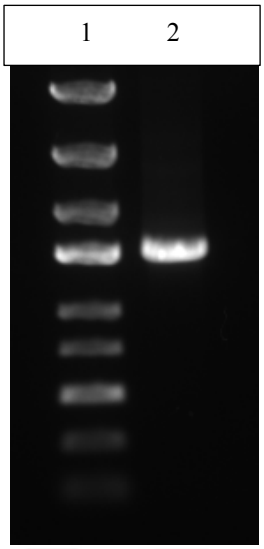


Figure 4. 12 Agarose gel electrophoresis of purified double digested Hygromycin with TEF promoter and TEF terminator fragment 1: GeneRuler Express; 2: Digested insert with *Bam*HI and *Pci*I REs after PCR purification.

Vector plasmid, pADH2-wt::*ADH2*, was also double digested with *Bam*HI and *Pci*I restriction enzymes. The whole plasmid is 4450 bp, and the region in which the REs cut out from the plasmid consisted of 1186 bp. After the double digestion, two separate fragments 1186 bp and 3264 bp were formed. (Figure 4.13)

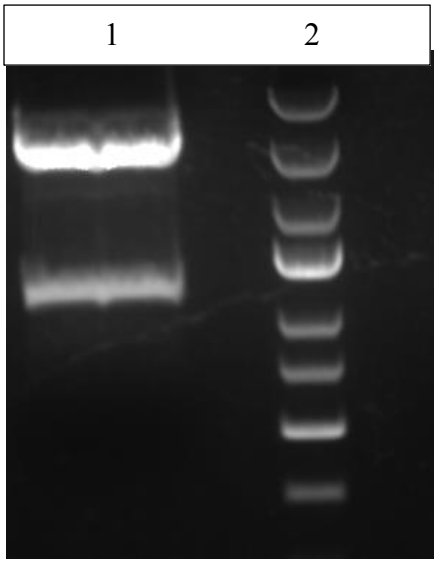


Figure 4. 13 Agarose gel electrophoresis image of 1: Double digested vector with *Bam*HI and *Pci*I; 2: GeneRuler Express.

Both double digested vector plasmid pADH2-wt::*ADH2* and the insert hygromycin fragment with its own TEF promoter and the terminator fragments, was extracted of from the gel by the razor procedure, purified with gel elution kit, and visualized in agarose gel electrophoresis after digestion (Figure 4.14). Double digested plasmid, pADH2-wt::*ADH2* is 3264 bp.

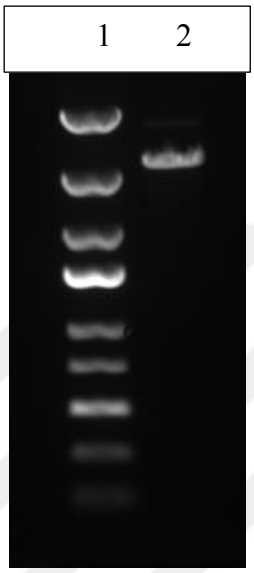


Figure 4. 14 Agarose gel electrophoresis of double digested vector-plasmid pADH2-wt::*ADH2* 1: GeneRuler Express; 2:Digested and purified vector-plasmid with *Bam*HI and *Pci*I.

The ligation reaction product was used in the transformation to *E. coli* DH5α cells by the calcium chloride method. To verify whether zeocin resistance gene was replaced with hygromycin resistance-gene, 12 single colonies from solid agar containing the selective antibiotic was inoculated into LB broth medium. After 16-18 hours of the incubation at 37°C, colony PCR with Forward_Hyg and Reverse_Hyg primers was performed. For colony PCR, thermocycling operation conditions and reaction mixture contents are presented in Tables 4.1 and 4.2. The PCR products visualized in agarose gel electrophoresis (Figure 4.15). Hygromycin resistance gene sequence with TEF promoter and TEF terminator is 1479 bp.

Forward_Hyg and Reverse_Hyg primers cover the same number of base pairs. The bands around 1500 bp are the verified positive results.

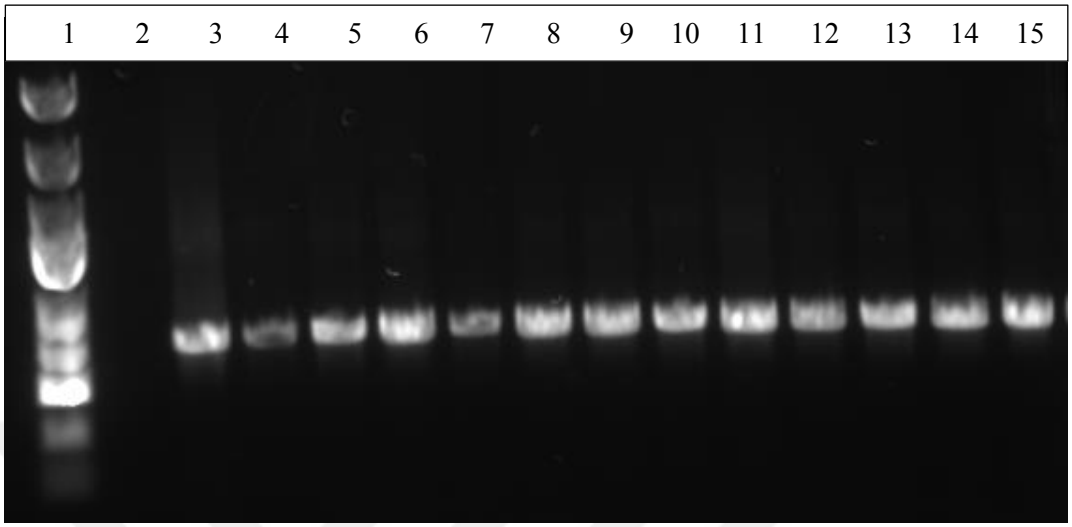


Figure 4. 15 Agarose gel electrophoresis of potential *E. coli* clones carrying the Hygromycin gene after transformation. 1: GeneRuler Express; 2: Negative control for PCR control of insert Hygromycin gene; 2: Negative control for PCR control of insert Hygromycin gene, 4-15: PCR products of potential colonies.

For the selection of the correct colonies, which carry the hygromycin-resistance gene but not the pADH2-wt::*ADH2* cassette, another PCR with the same colonies was performed using pADH2_forward and ADH2_reverse primers. The PCR products were visualized in agarose gel electrophoresis (Figure 4.16).

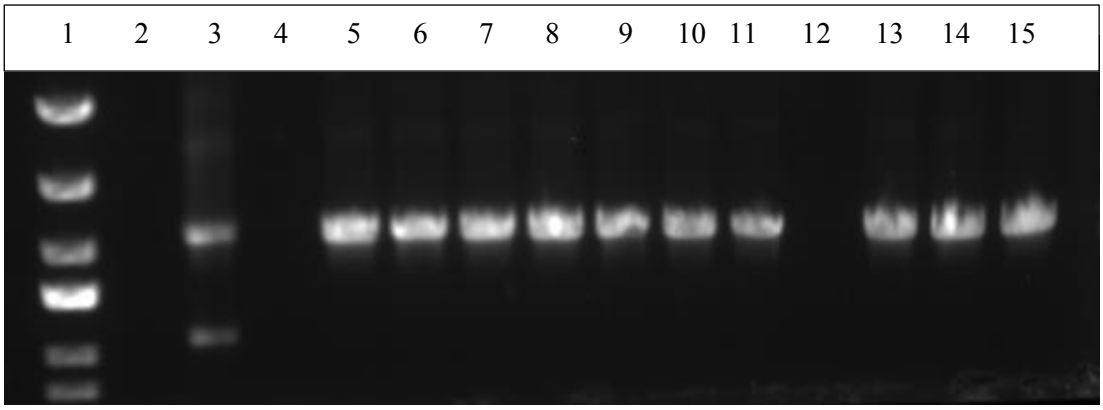


Figure 4. 16 Agarose gel electrophoresis of potential *E. coli* clones carrying the pADH2-wt::*ADH2* gene after transformation. 1: GeneRuler Express; 2: Negative *control for PCR control of insert* pADH2-wt*::ADH2 gene; 2: Negative control for PCR control of insert* pADH2-wt*::ADH2 gene, 4-15: PCR products of potential colonies.*

As the second step of the verification procedure, four colonies were selected, and putative plasmids were isolated, visualized in agarose gel electrophoresis (Figure 4.17). The size of the circular plasmids carrying pADH2-wt::*ADH2* cassette and hygromycin-resistance are the same and 4766 bp.

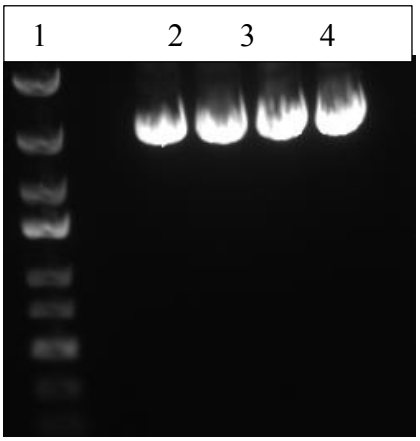


Figure 4. 17 Agarose gel electrophoresis of potential recombinant plasmids. 1: Generuler Express, 2-5: Potential plasmids carrying the Hygromycin and pADH2-wt::*ADH2* cassettes.

Using putative plasmids as templates, PCR was conducted using a) Forward_Hyg and Reverse_Hyg, and b) pADH2_forward and ADH2_reverse primers, separately, and products of PCR were visualized in agarose gel electrophoresis (Figure 4.18 and 4.19). All selected colonies gave positive results both for hygromycin gene (1479 bp) and pADH2-wt::*ADH2* (2100 bp). After colony PCR, two colonies were selected, and the constructs between *Pci*I and *Bam*HI restriction sites were verified with DNA sequencing analysis (Sentebio, Ankara) (Appendix B). Glycerol stocks of *E. coli* strains that carry the verified constructed plasmid, pADH2-wt::*ADH2*, were prepared and stored at -80°C.

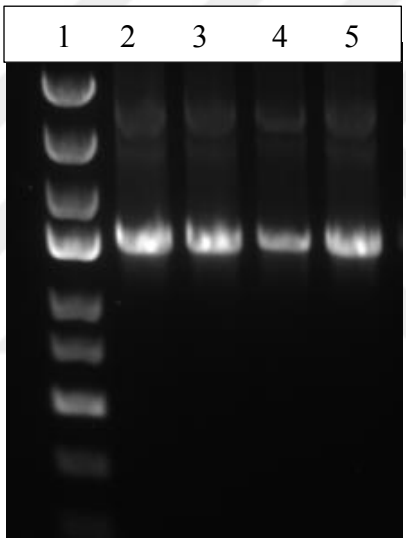


Figure 4. 18 Agarose gel electrophoresis of control PCR for potential recombinant plasmids. 1:Generuler Express; 2-5: Control PCR with Forward_Hyg and Reverse_Hyg primers. (Hygromycin gene 1479bp)

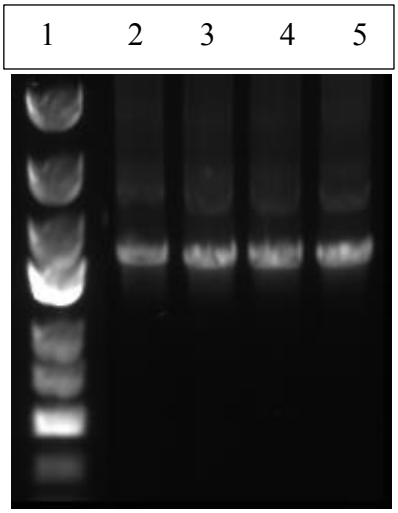


Figure 4. 19 Agarose gel electrophoresis of control PCR for potential recombinant plasmids. 1:Generuler Express; 2-5: Control PCR with pADH2_forward and ADH2_reverse primers. (pADH2-wt::*ADH2* cassette 2100 bp)

P. pastoris transformation

To transform the constructed plasmid into *P. pastoris* cells, pADH2-wt::*ADH2* plasmid was isolated from transformed *E. coli* DH5α strain The plasmid pADH2-wt*::ADH2* were linearised with *Bam*HI as indicated in Table 3.20, (Figure 4.20), and purified by Gel Extraction Kit (ThermoScientific, USA) prior to the transformation to *P. pastoris.* The constructed plasmids were integrated into AOX1 transcription termination locus to eliminate the effects of genomic integration sites. Linearized pADH2-wt::*ADH2* plasmid carrying hygromycin fragments is 4476 bp.

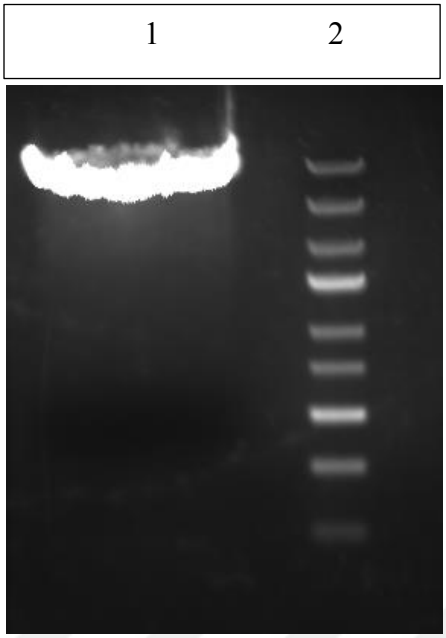


Figure 4. 20 Agarose gel electrophoresis of linearized plasmid pADH2-wt::*ADH2* plasmid carrying hygromycin fragments 1: pADH2-wt::*ADH2* plasmid after linearization with *Bam*HI ; 2: Generuler Express.

Linearized plasmid was transformed into wild type X-33 *P. pastoris* cells by lithium chloride method according to the procedure explained in section 3.4.3. After transformation, 25-100 µl of the cells were inoculated on a selective YPD agar medium containing 200 µg/ml hygromycin and incubated at 30°C for 48-60 hours. 16 colonies were selected from the proliferating cells, and was checked whether the transferred gene was integrated with the PCR. According to the colony PCR results, at least eight clones from each strain which was confirmed to carry the relevant gene were selected and inoculated into selective YPD media separately and incubated for 12-16 hours in air-filtered shake bioreactors at 200 rpm at 30°C. A glycerol stock was prepared from the proliferating recombinant cells to be used in future studies and stored at -80°C. Schematic representation of constructed *P. pastoris* expression system was shown in Figure 4.21.

Figure 4. 21 Schematic representation of constructed *P. pastoris* expression system.

4.2.1.2 Influences of *ADH2* overexpression on the growth of *P. pastoris* strains on the carbon sources ethanol and glucose

The influence of *ADH2* overexpression on the growth of *P. pastoris* strains constructed with $P_{ADH2\text{-}wt}$:: *ADH2* was investigated in batch cultivations of the carbon sources ethanol and glucose, separately. The carbon sources affected the cell growth as expected since, upon introduction to the central carbon metabolism, ethanol and glucose configure the metabolic reaction network architecture distinctly. The variations in the cell concentration of wild-type strain (*P. pastoris* X-33) and the concentrations of recombinant strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* overexpression cassettes with the cultivation time were determined in batch the cultivations on the carbon sources ethanol and glucose at concentrations of 2 % (*v/v*), and excess glucose (20 g/L). The growth profiles of the two colonies of the constructed strain with different copies of ADH2 and wild-type X-33 are presented in Figure 4.22.

On each substrate, the cell concentrations are close. The distinct loci of the variations in the cell concentrations with the cultivation time on glucose and ethanol demonstrate the dynamics of the growth and dependence on the carbon source (Figure 4.22). The results on glucose and ethanol reveal that the *P. pastoris* X-33 profiles are geometrically the envelope-curve of the each family-of-curves of the novel strains constructed. The fermentation of ethanol prolonged the lag-phases compared to that on glucose, and proceed through the exponential phase with a distinct loci.

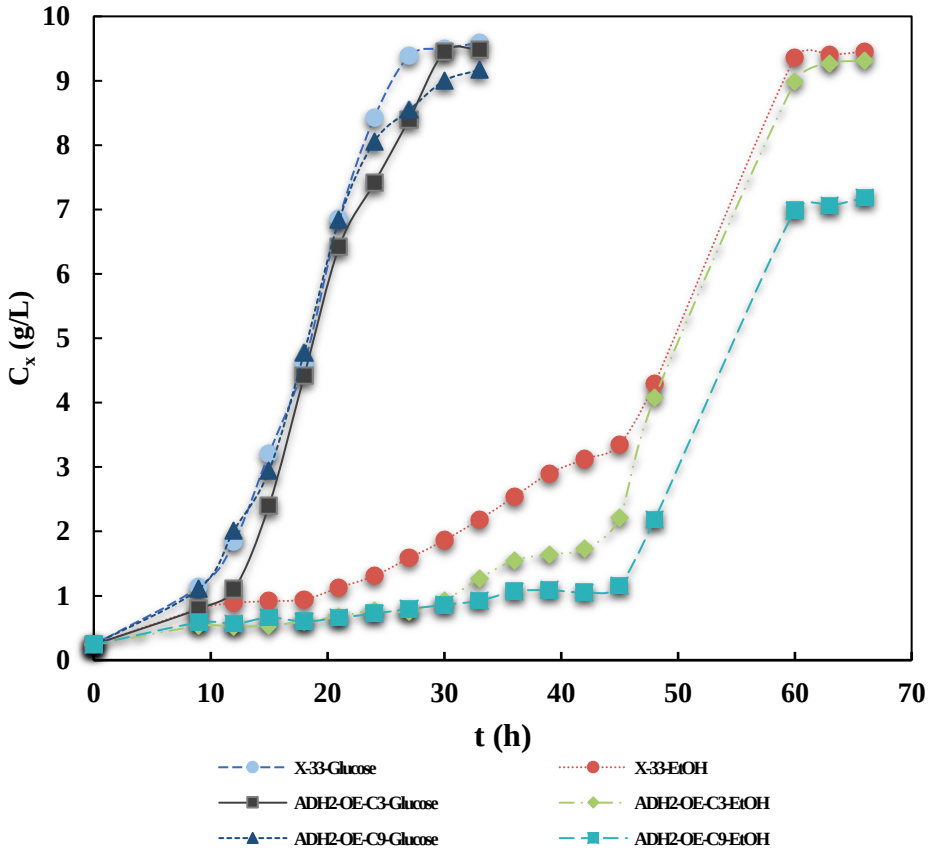


Figure 4. 22 Variations in the cell concentrations (DCW) of the strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* overexpression cassettes in batch fermentation of 2 % (v/v) ethanol, and excess glucose (20 g/L) and for comparison with *P. pastoris* X-33, with the cultivation time in shake-bioreactors in V_R = 50-mL fermentation volume at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}.

The results on glucose reveal that *P. pastoris* X-33 and recombinant strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* overexpression follow similar loci throughout the fermentation process. The duration of lag and exponential phases are nearly equal on glucose. When ethanol is used as carbon source, 1.35-fold higher cell concentrations were obtained in X-33 compared with *ADH2* overexpressed strains. Unlike glucose, the effect of overexpression of *ADH2* on the growth of recombinant strains is striking on ethanol so that reduced cell concentrations were obtained in all the overexpression strains during the first 45 hours of the fermentation. The cell growth is rather low in the overexpression strains. The low cell concentration in the first 45 hours and the growth profiles with different locus of the two colonies of the constructed strain are

related with the copy number of the *ADH2* gene in overexpression strains (Figure 4.22). Colonies carrying three *ADH2* gene copies were designated with C3 and colonies carrying nine *ADH2* gene copies were designated with C9. Thus, overexpression of *ADH2* results in growth inhibition. There is also an inverse correlation between the *ADH2* gene copy number in the overexpression strain and the cell growth.

With *ADH2* overexpression, it was aimed to increase the reaction fluxes in the ethanol utilization pathway and thus accelerate the ethanol metabolism. However, *ADH2* overexpression resulted in inhibition on the growth. Extracellular acetate concentrations were measured to investigate whether the cell growth inhibition, can be associated with the acetate response or not. At t = 36 h, the concentrations of extracellular acetic acid, which is the by-product of the EUT pathway, are presented in Figure 4.23.

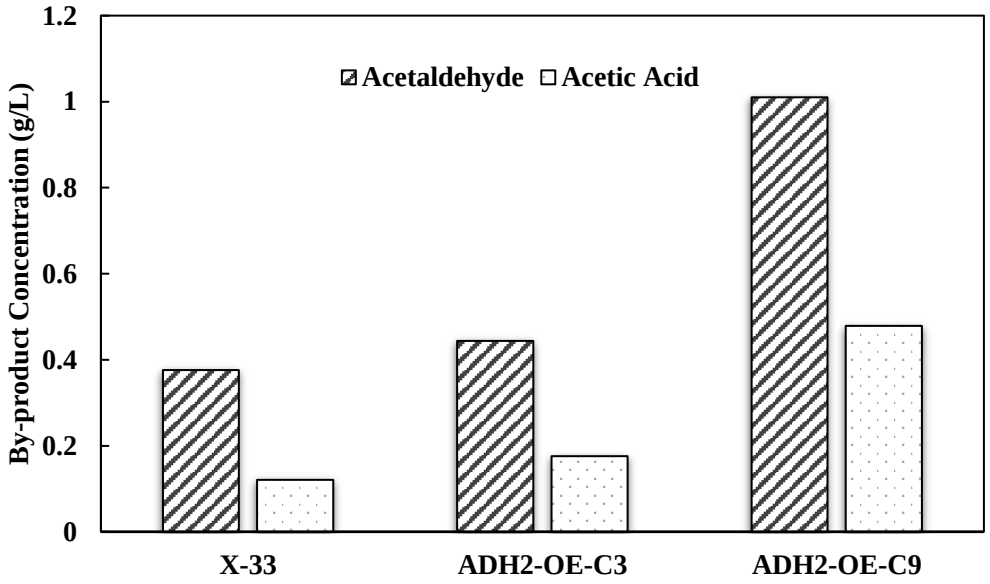


Figure 4. 23 By-product concentrations in the cell culture at t = 36 h for *ADH2* overexpression strains and wild-type *P. pastoris* X-33 strain on 2% (v/v) ethanol

The results presented in Figure 4.23 reveal that overexpression of the *ADH2* gene resulted in the accumulation of acetic acid and acetaldehyde in the cell. There is a direct correlation between *ADH2* gene copy number and by-product accumulation. While acetic acid accumulation in the C3 reached 0.2 g/L, acetic acid accumulation in the C9 reached up to 0.5 g/L, and acetaldehyde accumulation in C3 and C9 was reached 0.44 g/L and 1.01 g/L, respectively. While accumulation of acetaldehyde and acetic acid with wild-type cell were ca. 0.37 and 0.12 g/L, respectively. Using ethanol at initial concentration of 2% (v/v) caused higher acetaldehyde accumulation then the acetic acid accumulation due to the pressure caused by high ethanol concentration. For *ADH2*-C3 and *ADH2*-C9, acetaldehyde concentrations are 2.52-fold and 2.11-fold higher than acetic acid concentrations.

Influences of *ADH2* overexpression on the growth of *P. pastoris* on ethanol and glucose have been studied for the first time. Maestre *et al.* (2008) reported that overexpression of ADH2 enzyme in *Saccharomyces bayanus* caused a decrease in cell concentration after 48 h of the fermentation in YNB medium, similar to our results. One of the possible causes of this phenomenon is due to the accumulation of acetic acid, acetaldehyde, and NADH caused by overexpression of *ADH2*, the deterioration of the redox balance, consequently accumulation of toxic substances. This is in agreement with the fact that weak acids can inhibit the cell growth, cytostasis or cell death depending on the amounts and lipophilic moiety (Stratford and Anslow, 1996, 1998; Piper *et al.,* 2001), especially acetic acid has recently been shown to cause yeast cell death and aging (Giannattasio *et al.,* 2013). Simpson-Lavy and Kupiec (2019) showed that changes in the ethanol-acetate pathway alter acetic acid synthesis during fermentation, and because acetate is hazardous to cells, results in cell death and a shorter life span.

Inan and Meagher (2001) also stated that when ethanol is used as carbon source, the rate-limiting step appears to be the assimilation of acetate, resulting in acetate accumulation, which is in agreement with the results reported in this thesis.

4.2.2 Overexpression of Acetyl-CoA synthetase 1 (ACS1) enzyme

To overcome limitations in ethanol utilization of *P. pastoris* by increasing the expression of rate-limiting step; consequently, with the aim of increase ethanol uptake rate and regulatory role of ethanol as carbon source, *ACS1* overexpression was studied.

Construction of overexpression cassette of *ACS1* was designed under the control of pAOX1-v, which was constructed in our research group (Ergun *et al.* 2020). The primary reason for choosing this promoter variant was that it is inducible in the presence of ethanol. It is a tightly regulated novel engineered promoter variant, pAOX1/Cat8-L3, with moderate promoter strength. The reason for not using high strength promoters is that it may cause disruption of the cell metabolism by the over-accelerating overexpressed reaction. Naturally occurring promoters such as P_{ADH2} and P_{ACS1} was not used since the regulation of these promoters was directly affected by acetate concentration in the culture medium.

For selection based on kanamycin, pGAPZ-A-KanMX plasmid, was constructed by replacement of zeocin resistance gene with Kanamycin resistance gene, and used as vector plasmid. To be able to combine *ADH2* and *ACS1* overexpression cassettes by sequential transformation and to investigate the effects on the cell growth and ethanol uptake rate, a new antibiotic resistance gene (other than Hygromycin) was used.

4.2.2.1 Construction of the recombinant plasmid and strains carrying pAOX1/Cat8-L3::*ACS1*

Recombinant plasmid bearing the pAOX1/Cat8-L3::*ACS1* cassette was constructed. The recombinant plasmid was generated by inserting pAOX1/Cat8-L3 and *ACS1* into the vector-plasmid pGAPZα-A. *Kpn*I and *Bgl*II are the appropriate restriction enzymes, as in pADH2-wt::*ADH2* plasmid construction, since these enzymes are one-side cutter of the vector plasmid pGAPZ-A, and noncutter of pAOX1/Cat8-L3 and *ACS1* fragments. SnapGene and Restriction mapper was used to identify the

proper restriction sites. In Figure 4.24, vector plasmid pGAPZα-A and the restriction enzyme used in construction of recombinant plasmid are presented.

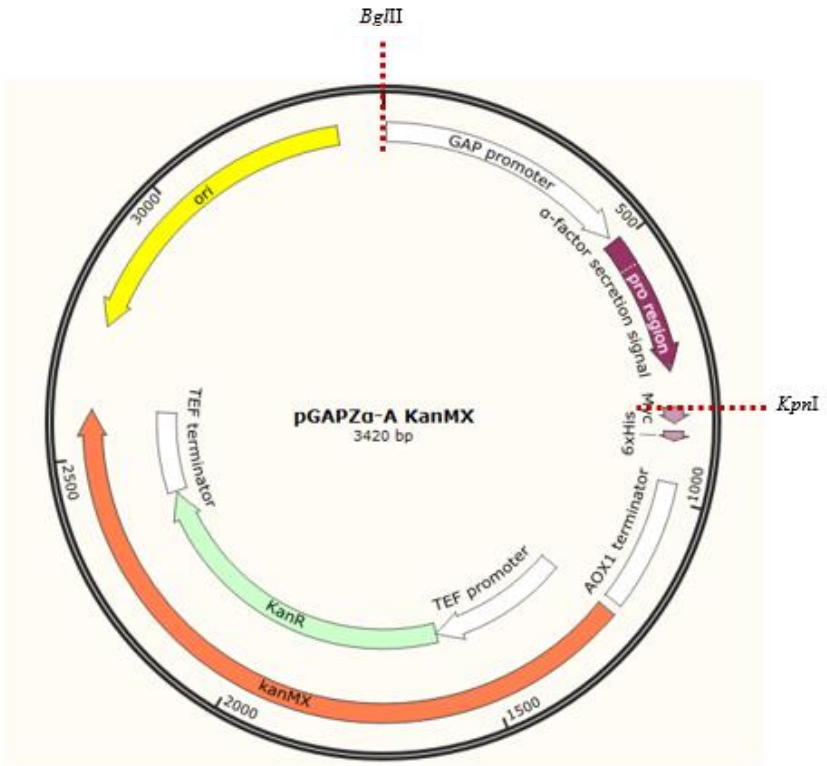


Figure 4. 24 Base plasmid pGAPZα-A (KanMX), and the restriction enzymes *Bgl*II and *Kpn*I, used in construction of recombinant plasmid.

The pGAPZ-A plasmid contains a *Bgl*II RE recognition site at the 5' end of pGAP; however, *Bgl*II truncates the pAOX1-Cat8-L3 gene, so the insert cannot be digested with *Bgl*II. *Bgl*II is required for digestion of the vector pGAPZ-A plasmid. As explained in section 4.2.1, *Bam*HI and *Bgl*II generate different but compatible ends, which makes the ligation possible. After ligation, however, the newly formed sequence can be recognized by neither *Bam*HI nor *Bgl*II. Thus, the *Bam*HI recognition region has been added to the pAOX1/Cat8-L3 5' end to be able to ligate the vector pGAPZα-A, and the insert pAOX1/Cat8-L3::*ACS1* (Figure 4.2). Since *Bam*HI cannot cut the newly formed restriction site, it is still a one-site cutter of the

plasmid and can be used for linearization of the plasmid before transformation to *P. pastoris* into AOX termination locus.

The pAOX1/Cat8-L3::*ACS1* was constructed using a two-step gene splicing by overlap extension (SOE) PCR. For the construction of pAOX1-Cat8-L3::*ACS1*, pAOX1-Cat8-L3 (940 bp) and *ACS1* (2007 bp) fragments were amplified by High-fidelity Q5 DNA Polymerase (NEB). The wild-type *P. pastoris* genome was isolated and used as a template for amplification of *ACS1*. For amplification of the pAOX1-Cat8-L3 fragment was isolated from pAOX1/Cat8-L3::*eGFP* plasmid (Ergun *et al.* 2020), and used as the template. To optimize the PCR reaction conditions, the effects of different annealing temperatures were investigated for pAOX1-Cat8-L3 and *ACS1* amplification. The PCR reaction compositions and operation conditions are given in Tables 3.14, 3.15, and 3.16. The length of pAOX1-Cat8-L3 and *ACS1* were 940 bp and 2007 bp, respectively (Figures 4.25 and 4.26).

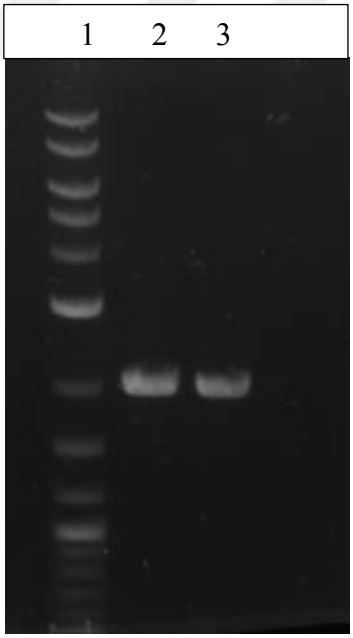


Figure 4. 25 Agarose gel electrophoresis of the gene sequences amplified with Forward_ACS1and Reverse_ACS1 primers at different annealing temperatures. 1: Quick-Load Purple 1 kb DNA ladder; Annealing temperatures: 2:T= 64.3°C; 3: T= 66.5° C.

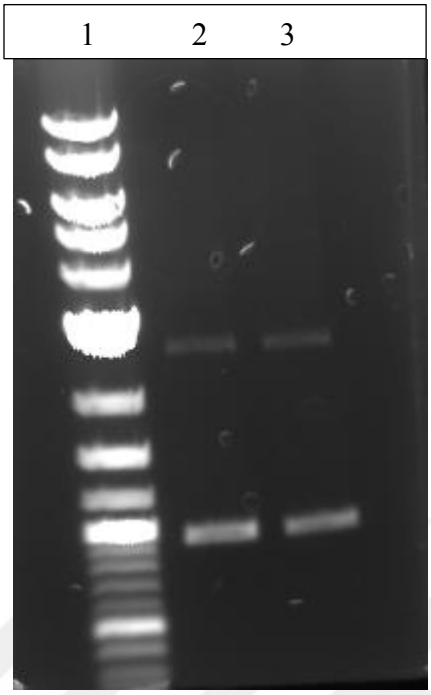


Figure 4. 26 Agarose gel electrophoresis image of the gene sequences amplified with Forward_pAOX1-Cat8-L3and Reverse_pAOX1-Cat8-L3 primers at different annealing temperatures. 1: G Quick-Load Purple 1 kb DNA ladder; Annealing temperatures: 2: T= 64.3° C; 3: T= 66.5° C.

The concentrations from the amplification of the pAOX-Cat8-L3 and *ACS1* gene fragment was similar for both annealing temperatures, 64.3°C and 66.5°C (Figure 4.26). The highest *ACS1* fragment concentration was obtained at the annealing temperature of T = 64.3°C (Figure 4.25). Thus, T = 64.3°C was determined as the optimum annealing temperature for the amplification of pAOX-Cat8-L3 and *ACS1* fragments; and for further amplification, was used as the optimum annealing temperature. The amplified fragments of pAOX-Cat8-L3 and *ACS1* was purified with GeneJet Gel Extraction Kit (ThermoFisher, USA) since other DNA fragments other than pAOX-Cat8-L3 and *ACS1* fragments–were also amplified. To obtain purified pAOX-Cat8-L3 and *ACS1* fragments, gel extraction was conducted, visualized in agarose gel electrophoresis after purification. (Figure 4.27)

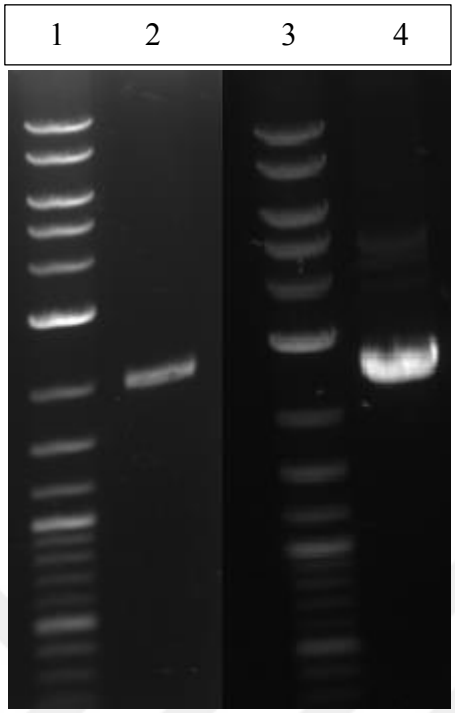


Figure 4. 27 Agarose gel electrophoresis image of the purified DNA fragments of pAOX1-Cat8-L3 and *ACS1*. 1&3: Quick-Load Purple 1 kb DNA ladder; 2: Purified DNA fragment of *ACS1* 4: Purified DNA fragment of pAOX1-Cat8-L3.

After purification of amplicons with overlapping compatible ends, the fragments were used a templates for the SOE-PCR. Amplicon pairs purified in the first step of the SOE-PCR were coupled with the primers Forward_pAOX-Cat8-L3 and Reverse_ACS1 with the PCR mixture compositions and thermocycling conditions given in Table 3.17 and 3.18, respectively. pAOX-Cat8-L3 and *ACS1* fragments fused for pAOX-Cat8-L3::*ACS1* synthesis.

Because *Kpn*I and *Bgl*II cannot operate in a single buffer, the base-plasmid pGAPZα-A was digested sequentially with the restriction enzymes *Kpn*I and *Bgl*II, respectively. Following the first restriction reaction, PCR purification was carried out. *Bam*HI and *Kpn*I, can also work in a single buffer, so insert, pAOX-Cat8-L3::*ACS1* was double digested with *Bam*HI and *Kpn*I in a single digestion reaction, and visualized in agarose gel electrophoresis. (Figure 4.28)

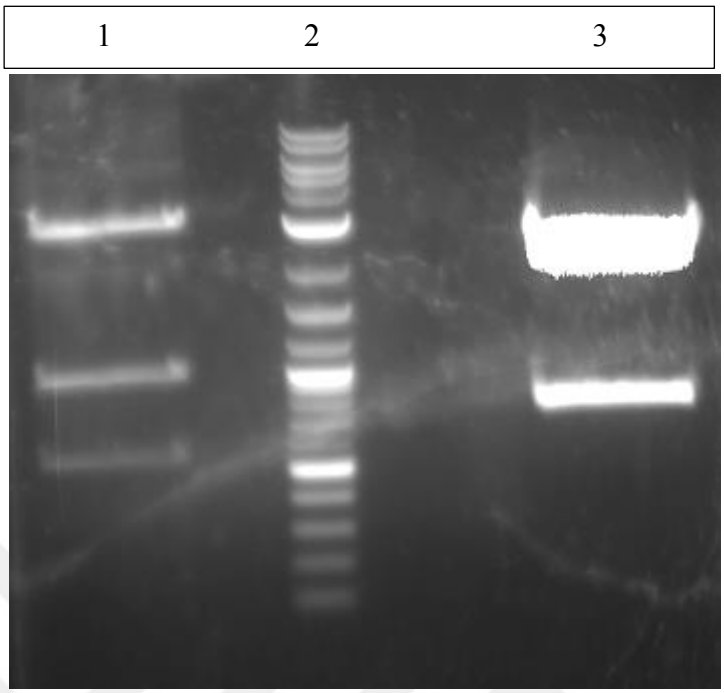


Figure 4. 28 Agarose gel electrophoresis of double digested vector, pGAPZA , and insert, pAOX-Cat8-L3::*ACS1*. 1: Digested insert with *Bam*HI and *Kpn*I; 2: Quick-Load Purple 1 kb DNA ladder; 3: Digested base-plasmid, pGAPZαA-KanMX, with *Kpn*I and *Bgl*II. (Before gel extraction)

After completing the second digestion reaction of the base-plasmid and the double digestion reaction of the insert, the DNA fragments were extracted from the gel using the razor procedure, purified with the gel elution kit, visualized in agarose gel electrophoresis. (Figure 4.29)

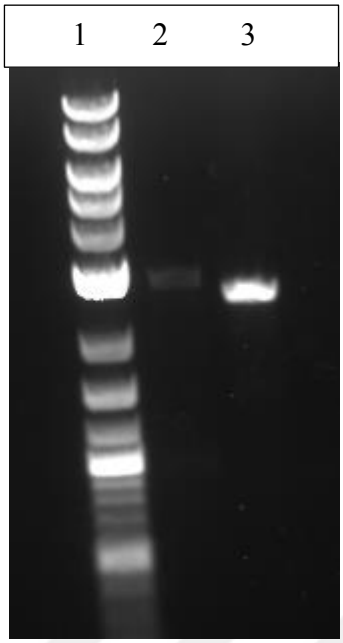


Figure 4. 29 Agarose gel electrophoresis of purified double digested vector, pGAPZαA, and insert, pAOX-Cat8-L3::*ACS1*.1: Quick-Load Purple 1 kb DNA ladder; 2: Digested insert with *Bam*HI and *Kpn*I REs after gel elution; 3: Digested base-plasmid, pGAPZαA-KanmMX, with *Kpn*I and *Bgl*II REs after gel elution.

Under the circumstances stated in Section 3.4.1.4, the ligation reaction was performed to clone the double digested pAOX1-Cat8-L3::*ACS1* into double digested vector pGAPZαA, using the vector: insert 3:1 ratio.

Transformation of the generated plasmid into wild-type *E. coli* DH5 cells was done after the ligation reaction using the calcium chloride technique. To make the selection, inoculation was done onto agar medium that contains a kanamycin-containing agar media. Then, a single colony from the solid agar containing the selective antibiotic was inoculated into LB broth medium. Twelve colonies were chosen among the probable transformants after 16-18 hours of inoculation at 37°C. Colony PCR was performed using Forward pAOX1-Cat8-L3 and Reverse *ACS1* primers to control the integration of plasmid was achieved or not, and visualized in agarose gel electrophoresis after colony PCR (Figure 4.30). pAOX1-Cat8-L3::*ACS1* is 2947 bp. For the control PCR, Taq DNA Polymerase (Thermo Scientific) was used. The colony PCR-operating parameters and reaction mixture composition are listed in the previous section, in Tables 4.1 and 4.2, respectively.

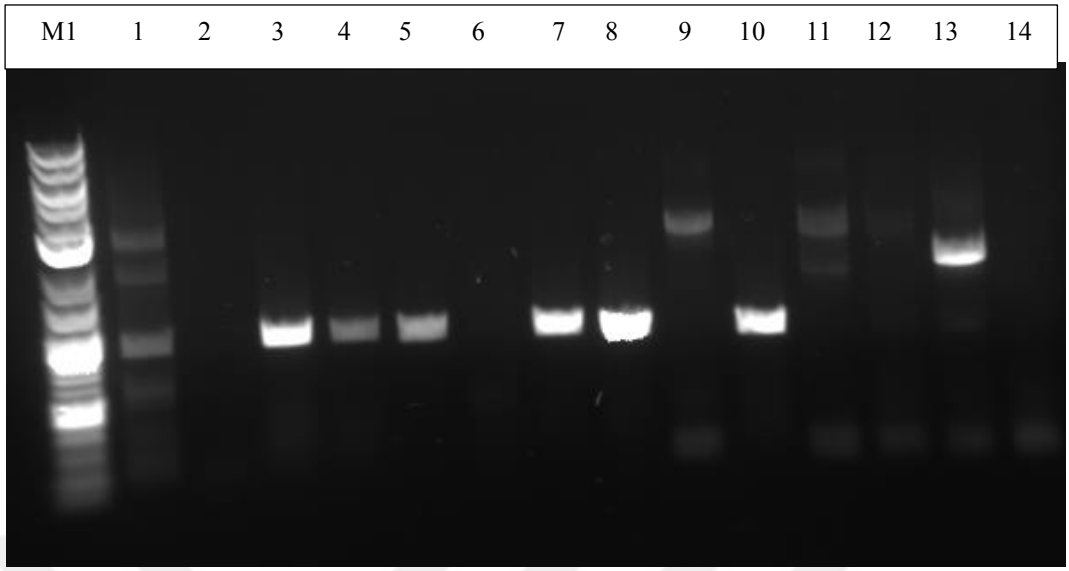


Figure 4. 30 Agarose gel electrophoresis image of potential *E. coli* clones carrying the $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* gene after transformation. M1: Quick-Load Purple 1 kb DNA ladder; 1: Positive control for PCR control of insert $P_{AOX\text{-}Cat8\text{-}L3}$-*ACS1* gene; 2: Negative control for PCR control of insert $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* gene, 3-14: PCR products of potential plasmids

Four colonies were selected among the positive colonies. Plasmid isolation was performed on four positive presumptive transformants to confirm plasmid integration, visualized after colony PCR. (Figure 4.31)

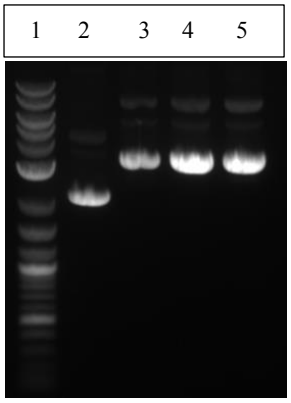


*Figure 4. 31*Agarose gel electrophoresis image of potential recombinant plasmids. 1: Quick-Load Purple 1 kb DNA ladder, 2-5: Potential plasmids carrying the $P_{AOX\text{-}}$

Cat8-L3::*ACS1* Another colony PCR was performed with Forward_pAOX1Cat8-L3 and Reverse_ACS1 primers under the conditions described in Tables4.1and 4.2, using isolated plasmids as a template. (Figure 4.32)

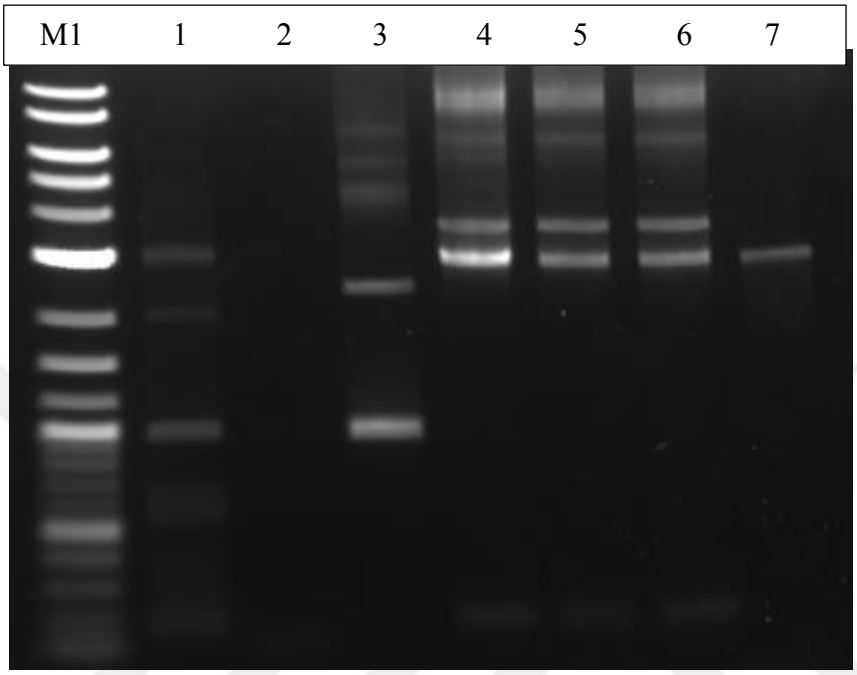


Figure 4. 32 Agarose gel electrophoresis of potential *E. coli* clones carrying the pAOX1-Cat8-L3::*ACS1* gene after transformation. M1: Quick-Load Purple 1 kb DNA ladder; 1: Positive control for PCR control of insert pAOX-Cat8-L3-*ACS1* gene; 2: Negative control for PCR control of insert pAOX-Cat8-L3-*ACS1* gene, 3-7: PCR products of potential plasmids

After the colony PCR, two colonies were selected and the constructs between *Bgl*II and *Kpn*I restriction sites were verified with DNA sequencing analysis (Sentebio, Ankara) (Appendix B). Glycerol stocks of *E. coli* strains that carry the verified constructed plasmid, pAOX1-Cat8-L3::*ACS1*, were prepared and stored at -80°C.

Transformation to *P. pastoris*

First, plasmid isolation from an *E. coli* DH5 strain was performed to achieve transformation of the generated plasmid into *P. pastoris* cells. After that, plasmids carrying the $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* cassette were linearized with *Bam*HI as specified in Table 3.32 (Figure 4.33), and the linearized plasmid was extracted from the gel by razor procedure.

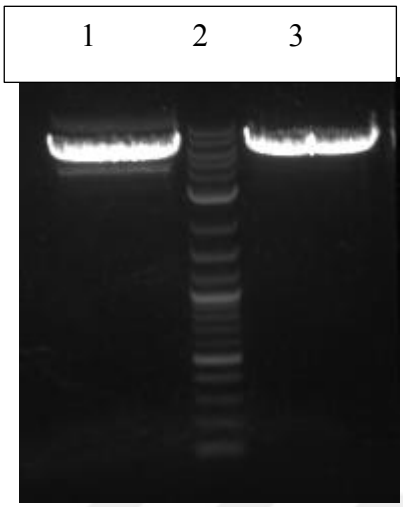


Figure 4. 33 Agarose gel electrophoresis image of 1&3: pAOX1-Cat8-L3::*ACS1* plasmid after linearization with *Bam*HI RE; 2: Quick-Load Purple 1 kb DNA ladder.

The linearized $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1* plasmid was transformed into wild type X-33 *P. pastoris*, using the lithium chloride technique, as detailed in section 3.4.3. After transformation, 25-100 µl of cells were inoculated onto a selective YPD agar medium which contains different amounts of G418 antibiotics in the ranch from 500 g/ml to 750 g/ml and incubated for 48-60 hours at 30°C. Following the incubation, 12 colonies for each transformation were chosen from the growing cells, and the transfected gene was verified for integration using the PCR. According to colony PCR results, at least eight clones from each strain that were confirmed to carry the necessary gene were selected and inoculated into selective YPD media separately and incubated for 12-16 hours in air-filtered shake bioreactors at 30°C and 200 rpm. For future research, a glycerol stock was made from the growing recombinant cells, preserved at -80°C. Figure 4.34 depicts a schematic illustration of the developed *P. pastoris* expression system.

Figure 4. 34 Schematic representation of the constructed *P. pastoris* expression systems.

4.2.2.2 Gene-copy number determinations of *ADH2*, *ACS1*, and *mApple* in selected *P. pastoris* Strains

Copy number of overexpressed genes is a critical parameter for correlation between overexpression and r-protein production strength of the constructed *P. pastoris* strains. Furthermore, it is also necessary for confirmation of the achievement of the construction of overexpression strains. The *ADH2* and *ACS1* gene copy numbers of heterologous genes were measured by quantitative polymerase chain reaction (qPCR) analysis (LightCycler® 480 Instrument II and LightCycler® 480 SYBR Green I Master, Roche, Basel, Switzerland) relative to the housekeeping gene *ARG4*. As a first step, the genomic DNA of the selected *P. pastoris* strains were isolated with Wizard Genomic DNA purification kit (Promega), visualized in agarose gel electrophoresis. (Figure 4.35 and 4.36)

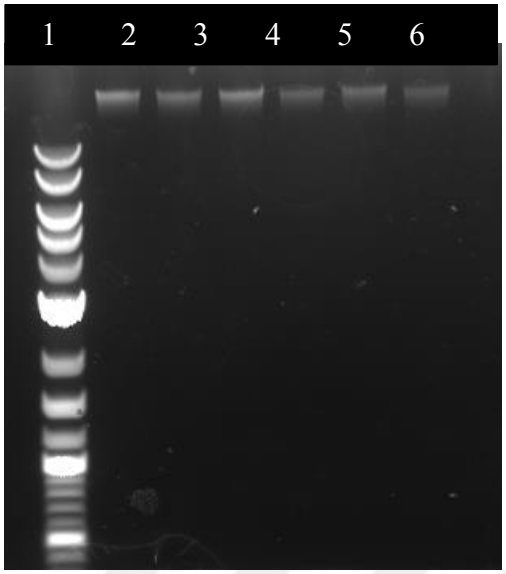


Figure 4. 35 Agarose gel electrophoresis of isolated genomic DNA of selected *P. pastoris* strains 1: Quick-Load Purple 1 kb DNA ladder; 2-7: Isolated Genomic DNA for selected colonies (600-1, 600-2, 600-5, 600-6, 750-1, 750-2) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1+* $P_{ADH2\text{-}Cat8\text{-}L2}$::*mApple* expression system.

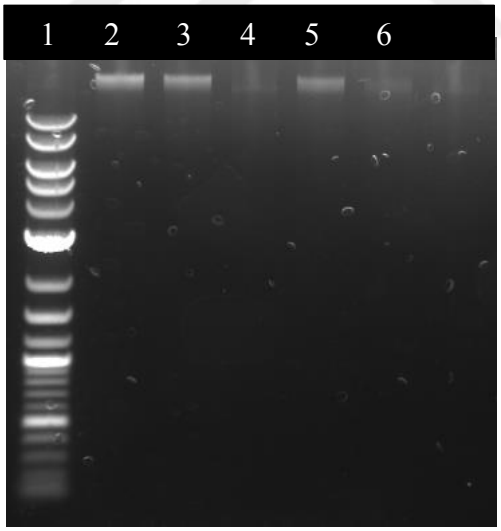


Figure 4. 36 Agarose gel electrophoresis of isolated genomic DNA of selected *P. pastoris* strains 1: Quick-Load Purple 1 kb DNA ladder; 2-3: Isolated Genomic DNA for selected colonies (12-13) of *P. pastoris* strains that contain $P_{ADH2\text{-}wt}$::*ADH2* plasmid; 5: Isolated Genomic DNA for the selected colony (500-1) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1*; 6: Isolated Genomic DNA for the selected colony (600-2) of *P. pastoris* strains that contain $P_{AOX\text{-}Cat8\text{-}L3}$::*ACS1+* $P_{ADH2\text{-}wt}$::*ADH2* expression system.

In addition, concentrations of isolated genomic samples were measured by nanodrop. (Table 4.3)

Table 4. 3 *Nanodrop measurement results of genomic DNA samples isolated from different P. pastoris strains*

***P. pastoris* Strain**	**Concentration (ng/µl)**	**260/280**	**260/230**
pADH2-wt::*ADH2*-C3	207.5	1.94	1.14
pADH2-wt::*ADH2*-C9	307.9	1.99	1.4
pAOX-Cat8-L3::*ACS1-C3*	90	1.92	0.72
pADH2-wt::*ADH2*-C3 + pAOX-Cat8-L3::*ACS1*-C3	26.9	1.69	0.31
pADH2-Cat8-L2::*mApple C1* + pAOX-Cat8-L3::*ACS1* C3	566.1	1.98	1.46
pADH2-Cat8-L2::*mApple C1* + pAOX-Cat8-L3::*ACS1* C4	411.7	1.95	1.41
pADH2-Cat8-L2::*mApple C1* + pAOX-Cat8-L3::*ACS1* C5	316.7	1.95	1.27
pmAOX::*mApple*	125	1.86	0.91
X-33	88.1	1.81	0.71

The procedure explained in detail in section 3.5 was followed to investigate the copy numbers of the genes. In this context, since wild type *P .pastoris* contains just one *argininosuccinate* gene (*ARG4*) in its genome, this gene was selected as the housekeeping gene, and quantification of examined genes was conducted relative to *ARG4*. In order to prepare standard calibration curves for accurate quantification

with the qPCR, *ARG4*, *ADH2*, and *ACS1* genes were amplified from the isolated *P. pastoris* wild-type strain with their standard primers (Table 3.34). Thermo-cyclic conditions and reaction compositions used for amplification of standards are also given in Table 3.35 and 3.36. Furthermore, the *mApple* standard was amplified from a pADH2-Cat8-L2::mApple plasmid with its forward and reverse primers. After amplification, standards were purified, visualized in agarose gel electrophoresis. (Figure 4.37 and 4.38) Their concentrations were also measured by nanodrop before the qPCR experiment (Table 4.4).

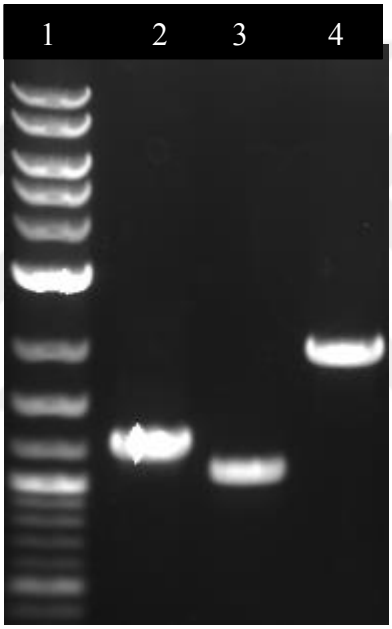


Figure 4. 37 Agarose gel electrophoresis of *ARG4, ADH2* and *ACS1* genes amplified by PCR 1: Quick-Load Purple 1 kb DNA ladder; 2: Amplified standard *ARG4* gene; 3: Amplified standard *ADH2* gene; 4: Amplified standard *ACS1* gene.

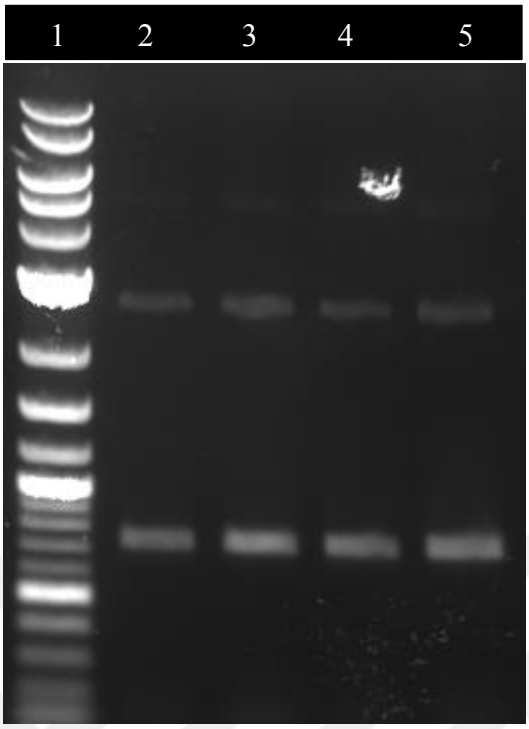


Figure 4. 38 Agarose gel electrophoresis of mApple gene amplified by PCR 1: Quick-Load Purple 1 kb DNA ladder; 2-5: Amplified *mApple* gene as standard.

Table 4. 4 *Nanodrop results of standards*

Standard	**Concentration (ng/µl)**	**Purity Ratio 260/280**	**Purity Ratio 260/230**
ARG4	339.4	1.84	1.43
ADH2	211.2	1.84	1.57
ACS1	215	1.83	1.55
mApple	46	1.79	1.03

After completing the concentration measurements with the nanodrop, gene copy numbers of the standards were calculated by the following formula in which 6.022*1023 and 109 represent the Avagadro number and conversion factor from Dalton to nanogram, respectively.

$$copy\ number = \frac{6.022 \times 10^{23} \times amount\ of\ sample}{Molecular\ weight\ of\ DNA \times 10^{9}} \quad (4.1)$$

Copy quantities of standard genes were calculated as fallows:

Copy-quantity of *ARG4* Standard = (339.4*6.022*10^{23})/ (777867.71*10^{9}) = 2.627*10^{11}

Copy-quantity of *ADH2* Standard = (211.2*6.022*10^{23})/ (650681.25*10^{9}) = 1.955*10^{11}

Copy-quantity of *ACS1* Standard = (215*6.022*10^{23})/ (1240045.15*10^{9}) = 1.044*10^{11}

Copy-quantity of *mApple* Standard = (46*6.022*10^{23})/(439436.87*10^{9}) = 0.630*10^{11}

Calibration curves of standards were obtained by preparing different dilutions that contain different copy-quantities (10^{3},10^{4},10^{5},10^{6}) and dilute them to 2 ng/ µl to use in the qPCR reaction mixture as a template and finally performed the qPCR reaction with LightCycler® 480 Instrument II (Roche). Experimental qPCR results of standard calibration curves were tabulated in Table 4.5 to 4.9.

Table 4. 5 *Standard curve results of ARG4*

Gene	**Cp**	**Concentration**	**Standard**
Arg 10^{3}	28.07	2.63E+03	2.63E+03
Arg 10^{4}	23.49	2.68E+04	2.63E+04
Arg 10^{5}	20.12	2.53E+05	2.63E+05
Arg 10^{6}	16.59	2.68E+06	2.63E+06

Table 4. 6 *Standard curve results of ADH2*

Gene	Cp	Concentration	Standard
ADH2 10^3	30.94	1.95E+03	1.95E+03
ADH2 10^4	24.57	2.14E+04	1.95E+04
ADH2 10^5	21.57	1.65E+05	1.95E+05
ADH2 10^6	17.83	2.10E+06	1.95E+06

Table 4. 7 *Standard curve results of ACS1*

Gene	Cp	Concentration	Standard
ACS1 10^3	35.78	1.04E+03	1.04E+03
ACS1 10^4	30.72	9.42E+03	1.04E+04
ACS1 10^5	25.77	1.31E+05	1.04E+05
ACS1 10^6	22.15	9.08E+05	1.04E+06

Table 4. 8 *Standard curve results of mApple*

Gene	Cp	Concentration	Standard
mApple 10^3	40	6.30E+02	6.30E+02
mApple 10^4	34.32	5.93E+03	6.30E+03
mApple 10^5	28.9	7.20E+04	6.30E+04
mApple 10^6	24.34	5.68E+05	6.30E+05

Table 4. 9 *Standard curve data of ARG4, ADH2, ACS1, and mApple*

	ARG4	***ADH2***	***ACS1***	***mApple***
Error	0.00492	0.0226	0.0313	0.0178
Efficiency	1.949	1.976	1.704	1.585
Slope	-3.450	-3.380	-4.322	-5
Y-intercept	38.77	39.21	47.90	52.18
Link	26800	21430	9418	5927

For qPCR, the specific primers of *ARG4, ADH2, ACS1*, and mApple were designed and used in the quantification of relative-gene-copy-numbers in the isolated colonies. Relative quantification results of qPCR were demonstrated in Table 4.10.

Table 4. 10 *Relative quantification results of selected P. pastoris strains.*

***P. pastoris* Strain**	**Quantified gene**	**Gene Copy Number**	**Target Cp**	**Reference Cp**
pADH2-wt::*ADH2* (C9)	*ADH2*	9.541	20.22	22.77
pADH2-wt::*ADH2* (C3)	*ADH2*	3.683	21.77	22.45
pAOX-Cat8-L3::*ACS1* (C3)	*ACS1*	4.017	23.14	21.69
pADH2-wt::*ADH2* + pAOX-Cat8-L3::*ACS1* (C3)	*ACS1*	3.613	23.84	22.14
pADH2-Cat8-L2::*mApple* + pAOX-Cat8-L3::*ACS1* (C3)	*ACS1*	3.775	24.65	22.2
pADH2-Cat8-L2::*mApple* + pAOX-Cat8-L3::*ACS1* (C4)	*ACS1*	4.595	23.43	22.66
pADH2-Cat8-L2::*mApple* + pAOX-Cat8-L3::*ACS1* (C5)	*ACS1*	5.38	22.9	22.47
pmAOX::*mApple*	*mApple*	1.216	28.06	22.24

For the constructed strains, the gene copy numbers of investigated genes *ADH2*, *ACS1*, and *mApple* were determined. The achievement of overexpression strains was

confirmed with different numbers of copies of constructed plasmids. Furthermore, gene copy number determination would allow to correlation production capacities of the constructed OE-*ACS1* expression systems with copy numbers.

In table 4.11, gene copy numbers of constructed *P. pastoris* strains and their abbreviations were listed.

Table 4. 11 *Constructed P. pastoris strains, gene copy numbers, and their abbreviations*

Constructed strains	**Gene copy number**	**Abbreviation**
$P_{ADH2\text{-}wt}$::*ADH2*	Three copies of *ADH2*	*ADH2*-OE-C3
$P_{ADH2\text{-}wt}$::*ADH2*	Nine copies of *ADH2*	*ADH2*-OE-C9
$P_{AOX1/Cat8\text{-}L3}$::*ACS1*	Three copies of ACS1	*ACS1*-OE-C3
$P_{ADH2\text{-}wt}$::*ADH2* + $P_{AOX1/Cat8\text{-}L3}$::*ACS1*	Three copies of *ADH2* Three copies of ACS1	*ADH2*-OE + ACS1-OE
$P_{AOX1/Cat8\text{-}L3}$::*ACS1* + $P_{ADH2\text{-}Cat8\text{-}L2}$::*mApple*	Four copies of *ACS1** One copy of *mApple*	*ACS1*-OE-C3+*mApple*-C1
$P_{AOX1/Cat8\text{-}L3}$::*ACS1* + $P_{ADH2\text{-}Cat8\text{-}L2}$::*mApple*	Five copies of *ACS1** One copy of *mApple*	*ACS1*-OE-C4+*mApple*-C1
$P_{AOX1/Cat8\text{-}L3}$::*ACS1* + $P_{ADH2\text{-}Cat8\text{-}L2}$::*mApple*	Six copies of *ACS1** One copy of *mApple*	*ACS1*-OE-C5+*mApple*-C1

**P. pastoris* contains one copy of *ACS1* gene in its genome under P_{ACS1}

4.2.2.3 Influences of *ACS1* overexpression on the growth of *P. pastoris* strains on carbon source ethanol

The effect of *ACS1* overexpression on the growth of *P. pastoris* strains constructed with $P_{AOX1/Cat8\text{-}L3}$:: *ACS1* was investigated in batch cultivations on 1 % (*v/v*) ethanol. The cell concentrations of the wild-type strain (*P. pastoris* X-33) and recombinant strains constructed with pADH2-wt::*ADH2* and pAOX1/Cat8-L3::*ACS1* were compared. The growth profiles of constructed strains, $P_{ADH2\text{-}wt}$::*ADH2*, $P_{AOX1/Cat8\text{-}L3}$:: *ACS1*, $P_{ADH2\text{-}wt}$:: *ADH2*+$P_{AOX1/Cat8\text{-}L3}$:: *ACS1*, $P_{ADH2/Cat8\text{-}L2}$*::mApple*, $P_{AOX1/Cat8\text{-}L3}$:: *ACS*+ $P_{ADH2/Cat8\text{-}L2}$*::mApple* and wild-type X-33 are presented in Figure 4.39.

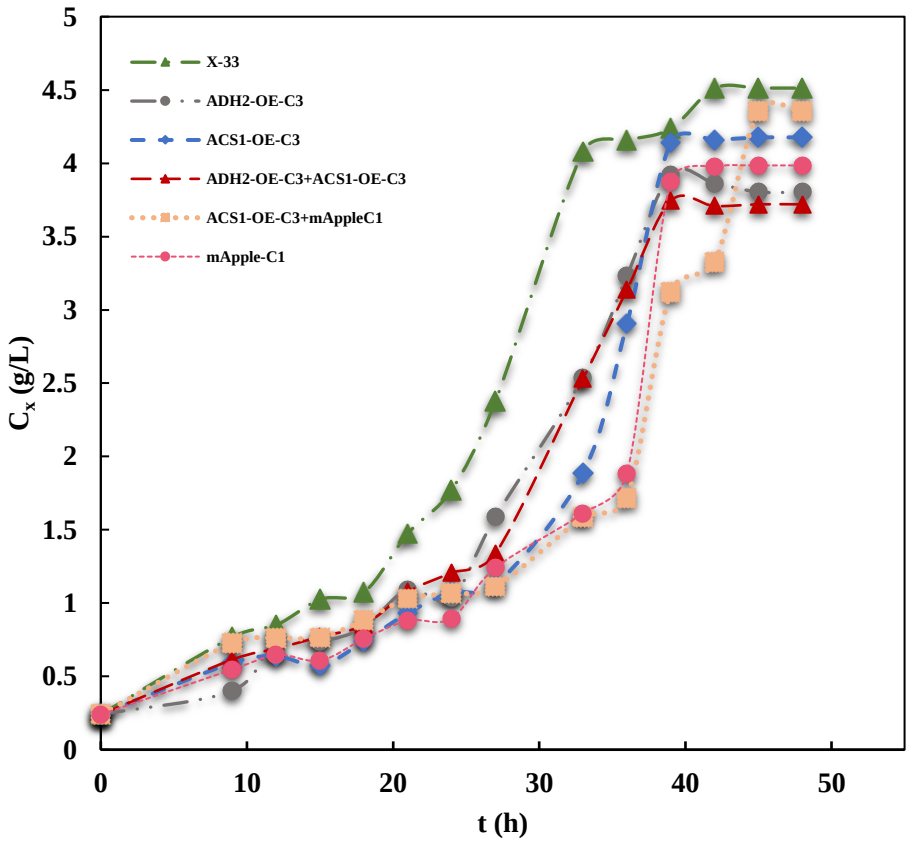


Figure 4. 39 Variations in the cell concentrations (DCW) of strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$::ACS1 overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33, with the cultivation time in shake-bioreactors in V_R = 50 mL fermentation volume at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}.

The close loci of the variations in the cell concentration with the cultivation time on ethanol demonstrate the dynamics of the growth and dependence on the carbon source (Figure 4.39). Starting from t = 0 and by diverging with a breakthrough at t = 15 h, the highest cell concentrations were obtained with P. pastoris x-33, which is geometrically the envelope of the family-of-curves of the novel strains constructed. The results on ethanol show that the recombinant strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$::*ACS1* overexpression follow similar trends for the first 24 h of the fermentation process. With the recombinant strains, the exponential phases started after t = 24 h with a characteristic breakthrough with the generated strains. The cells constructed with *ADH2*-OE and *ADH2*-OE+*ACS1*-OE, after t = 45 h reached the lowest cell concentrations, but with the highest slopes. Higher cell concentrations were obtained with *ACS1* overexpression strain than *ADH2* overexpression strain. Thus, it is concluded that the growth inhibition caused by *ADH2* overexpression is resolved by the *ACS1* overexpression since the *ACS1*-OE strain reached a cell concentration of 4.12 g/L, while the cell concentration was 3.7 g/L for the *ADH2* overexpression strain. Similar to that obtained with 2% (v/v) ethanol with *ADH2* overexpression strains, *ADH2*-OE strain produced lower cell concentration on 1% (v/v) ethanol compared to wild-type cell. However, on 2%(v/v) ethanol, the growth of the overexpressed *ADH2* strain was much slower than on 1% (v/v) ethanol. Wehbe, Yaman, and Çalık (2020) reported that although increasing ethanol concentrations results in an increase in the cell concentration at prolonged cultivation times, the cell growth on increased $C_{EtOH,0}$ was slower. In this context, it is possible to assert that increasing initial ethanol concentrations causes serious growth inhibition for *ADH2* overexpression compared to X-33. This is not due for the *ACS1* overexpression.

Variations in ethanol concentrations and specific ethanol uptake rates with the cultivation time are presented in Figures 4.40 and 4.41, respectively.

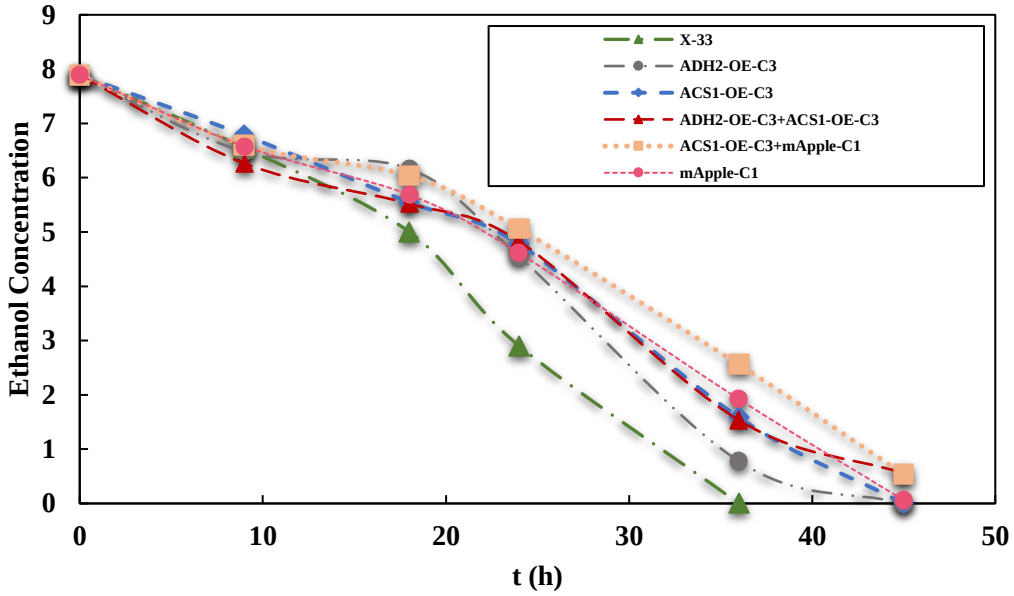


Figure 4. 40 Variations in ethanol concentrations of strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$:: *ACS1* overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33 with cultivation time

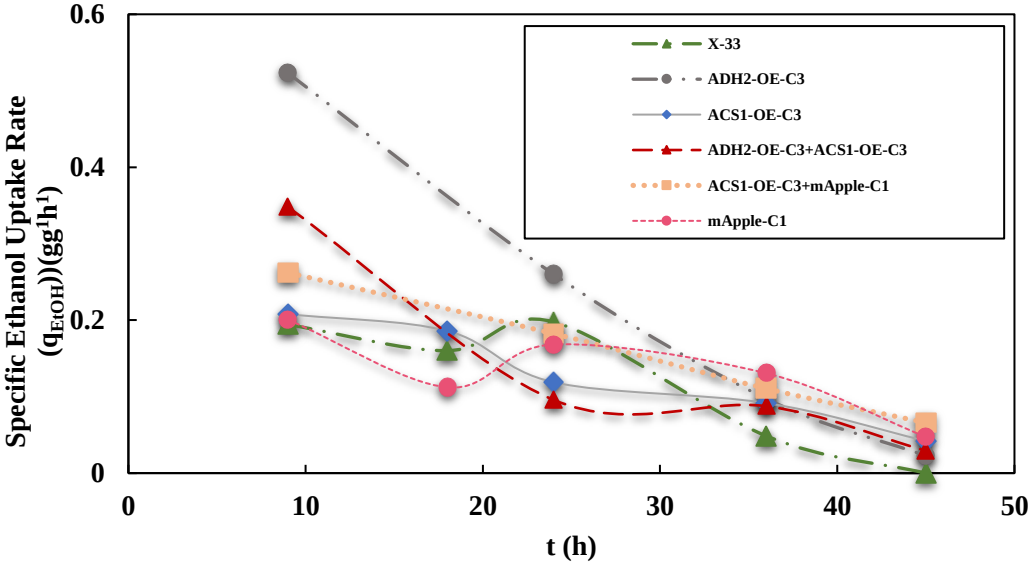


Figure 4. 41 Variations in the specific ethanol uptake rates of strains constructed with $P_{ADH2\text{-}wt}$::*ADH2* and $P_{AOX1/Cat8\text{-}L3}$:: *ACS1* overexpression cassettes in batch fermentation of 1 % (v/v) ethanol, and for comparison with *P. pastoris* X-33 with cultivation time

The highest specific ethanol uptake rate was obtained (q_{EtOH}) at t = 9 h for *ADH2-OE* and *ADH2-OE+ACS1-OE* strains as 0.52 and 0.35 g g^{-1} h^{-1}, respectively. Compared to wild-type X-33 (q_{EtOH} = 0.19 $gg^{-1}h^{-1}$), *ADH2* overexpression strain significantly increased the ethanol uptake rate. Although *ADH2-OE* increased the specific ethanol uptake rate (Figure 4.37), lower cell concentrations were obtained at the end of the fermentation. Compared to the wild-type strain (*P. pastoris* X-33), 1.1-fold and 1.22-fold lower cell concentrations were obtained with *ADH2*-OE and *ADH2-OE+ACS1-OE,* respectively. This implies that increasing specific ethanol uptake rate causes the formation of the by-products at a higher rate by the bottleneck consecutive reactions of the EUT pathway. This is in agreement with the fact that overexpressed *ADH2* strains result in growth inhibition due to acetate accumulation. On the other hand, variation in ethanol concentration during the first 9 hours was similar in all the constructed strains and X-33 strain. A significant decrease in the specific ethanol uptake rates was determined for the *ADH2* overexpression strains between $9 < t < 18$ h, while the decrease was with lower slopes for the *ACS1* overexpression, X-33, and pADH2/Ca8-L2::*mApple* strains. The ethanol concentrations until t = 18 h were profile except X-33. At t = 24 h the lowest specific ethanol uptake rate was obtained with the *ADH2-OE+ACS1-OE* strain. In addition, at $18 < t < 24$ h, X-33 and pADH2-Cat-L2::*mApple* strains showed similar specific ethanol uptake rate profiles and formed a small plateau, then, by t=24 h, started to decrease. For *ACS1* overexpression stains, the specific ethanol uptake rate was high in the first 18 hours, while low levels were obtained after t= 24 h. The decrease in the ethanol uptake can be because of the acetate accumulation which is less compared with *ADH2*-OE strain.

Acetaldehyde and acetic acid concentrations in fermentation broth at t=36 h are presented in Figure 4.32.

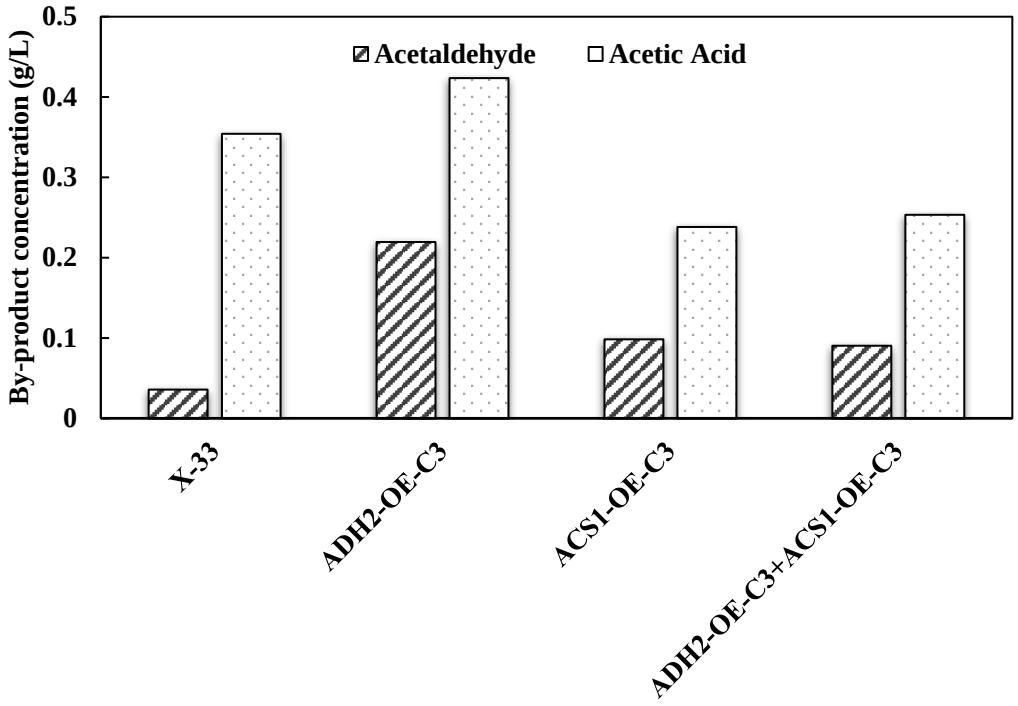


Figure 4. 42 By-product concentrations in the cell culture at t=36 h for overexpression strains and wild-type *P. pastoris* X-33 strain on 1% (v/v) ethanol

At t = 36 h of the fermentation, lower extracellular C_{AcAl} and C_{AcAc} were obtained with *ACS1*-OE strain as 0.23 g/L and 0.098 g/L, compared to *ADH2*-OE and *ADH2*-OE+*ACS1*-OE strains. With *ACS1*-OE strains, higher acetic acid accumulation was observed that is 2.3-fold higher the than acetaldehyde accumulation. This is in agreement with the fact that conversion of acetate to acetyl Co-A is the main rate-limiting step in the EUT pathway. The by-product concentration results are also consistent in the fact that overexpression of ACS1 decreases the acetate accumulation because overexpression of *ACS1* decreases acetic acid accumulation 1.52-fold and 1.83-fold compared to wild-type (*P. pastoris* X-33) and *ADH2* overexpression strains, respectively. Besides, 1.1-fold and 1.13-fold increase in the cell concentration was achieved with *ACS1*-OE compared to *ADH2*-OE and *ADH2*-

OE+*ACS1*-OE. There is no significant difference between *ACS1*-OE and *ADH2*-OE+*ACS1*-OE in terms of the by-product reduction. However, with *ACS1*-OE compared to *ADH2*-OE+*ACS1*-OE , 1.13-fold higher cell concentration was obtained; thus, *ACS1-OE* was selected as the superior strain for the r-protein production investigation studies.

4.3 Transcript level analysis of selected central carbon metabolism genes in the generated recombinant strains

After investigating the effects of the engineered expressions of the genes in the ethanol pathway on the cell growth and ethanol uptake rate, the transcript levels of the 13 genes in metabolically engineered strains on ethanol were analyzed. The wild-type and overexpression strains were grown on ASMv6 medium with 1% (v/v) ethanol. The cells were collected when their OD values reach OD = 6-7. Total RNA isolation was performed with QIAGEN RNaesy Mini Kit by following the protocol described in section 3.9. RNA quantity and quality were determined using the NanoDropTM 1000 Spectrophotometer; the average RNA yield and purity values are presented in Table 4.12.

Table 4. 12 *Quality and Quantity of P. pastoris RNA Samples Prepared by RNaesy Mini Kit (Qiagen)*

Samples	Concentration (ng/µl)	OD_{260}/OD_{280}
pADH2-wt::*ADH2* (C-9)	78.24	2.18
pADH2-wt::*ADH2* (C-3)	102.28	2.06
pAOX1-Cat8-L3::*ACS1*	91.19	1.86
pAOX1-Cat8-L3::*ACS1* + **pADH2-wt::*ADH2***	110.55	1.75
X-33	98.80	2.07

After isolation of total RNA, cDNA was synthesized using oligo(dT)23 primers and First Strand cDNA Synthesis Kit (Roche Life Science, Germany) by METU Centre Laboratory according to manufacturer's instructions. The Real-time PCR reactions were performed with primers listed in Table 3.45. Changes in transcript levels in *P. pastoris* mutant strains were compared to wild-type controls using the threshold cycle (ΔΔCT) method, using Equations 4.1 - 4.3.

$$(\Delta CTE) = TE - HE \quad (4.1)$$

$$(\Delta CTC) = TC - HC \quad (4.2)$$

$$(\Delta\Delta CT) = (\Delta CTE - \Delta CTC) \quad (4.3)$$

$$\text{Expression fold change} = 2^{-(\Delta\Delta CT)} \quad (4.4)$$

TE and TC represent the C_T values of the tested and housekeeping genes in constructed strains. HE and HC represent the C_T values of the tested and housekeeping genes in the control strain. All expression signals were normalized relative to the expression of the actin gene *ACT1*.

To identify potential effects of overexpression of *ADH2* and *ACS1* genes on central carbon metabolism of *P. pastoris*, mRNA levels of 13 genes (*ADH2, ALD4, ACS1, MLS1, PYC2, PDA1, PGK1, CDC19, FBP1, CAT8-1, CAT8-2, MXR1,* and *DAS1*) in TCA cycle, gluconeogenesis, and ethanol utilization pathway were analyzed (Table 4.13). Representative illustrations of the TCA cycle, gluconeogenesis, and ethanol utilization pathway are presented in Figure 4.43.

Table 4. 13 *Genes involved in central carbon metabolism of P. pastoris*

Gene	Encoding Enzyme	Reaction catalyzed
ADH2	Alcohol dehydrogenase 2 (EC: 1.1.1.1)	$EtOH + NAD^+ \longrightarrow$ Asetaldehyde $H^+ + NADH$
ALD4	Aldehyde dehydrogenase 4 (EC: 1.2.1.5)	Asetaldehyde + H_2O + $NADP^+$ $\longrightarrow$ Acetate + $2H^+$+ NADPH
ACS1	Acetyl-coenzyme A synthetase 1 (EC: 6.2.1.1)	Acetate + CoA $\longrightarrow$ Acetyl-CoA
MLS1	Malate synthase 1 (EC: 2.3.3.9)	Acetyl-CoA + H_2O + Glyoxylate $\longrightarrow$ Malate + CoA
PYC2	Pyruvate carboxylase 2 (EC: 6.4.1.1)	Bicarbonate + Pyruvate Oxaloacetate + Phosphate
PDA1	Pyruvate dehydrogenase (EC: 1.2.4.1)	Pyruvate $\longrightarrow$ Acetyl-CoA
PGK1	Phosphoglycerate kinase (EC: 2.7.2.3)	(2*R*)-3-phosphoglycerate $\longrightarrow$ (2*R*)-3-phospho-glyceroyl phosphate
CDC19	Pyruvate kinase 1 (EC: 2.7.1.40)	
FBP1	Fructose-1,6-bisphosphatase (EC: 3.1.3.11)	Fructose 1,6-bisphosphate + H_2O $\longrightarrow$ Fructose 6-phosphate + phosphate

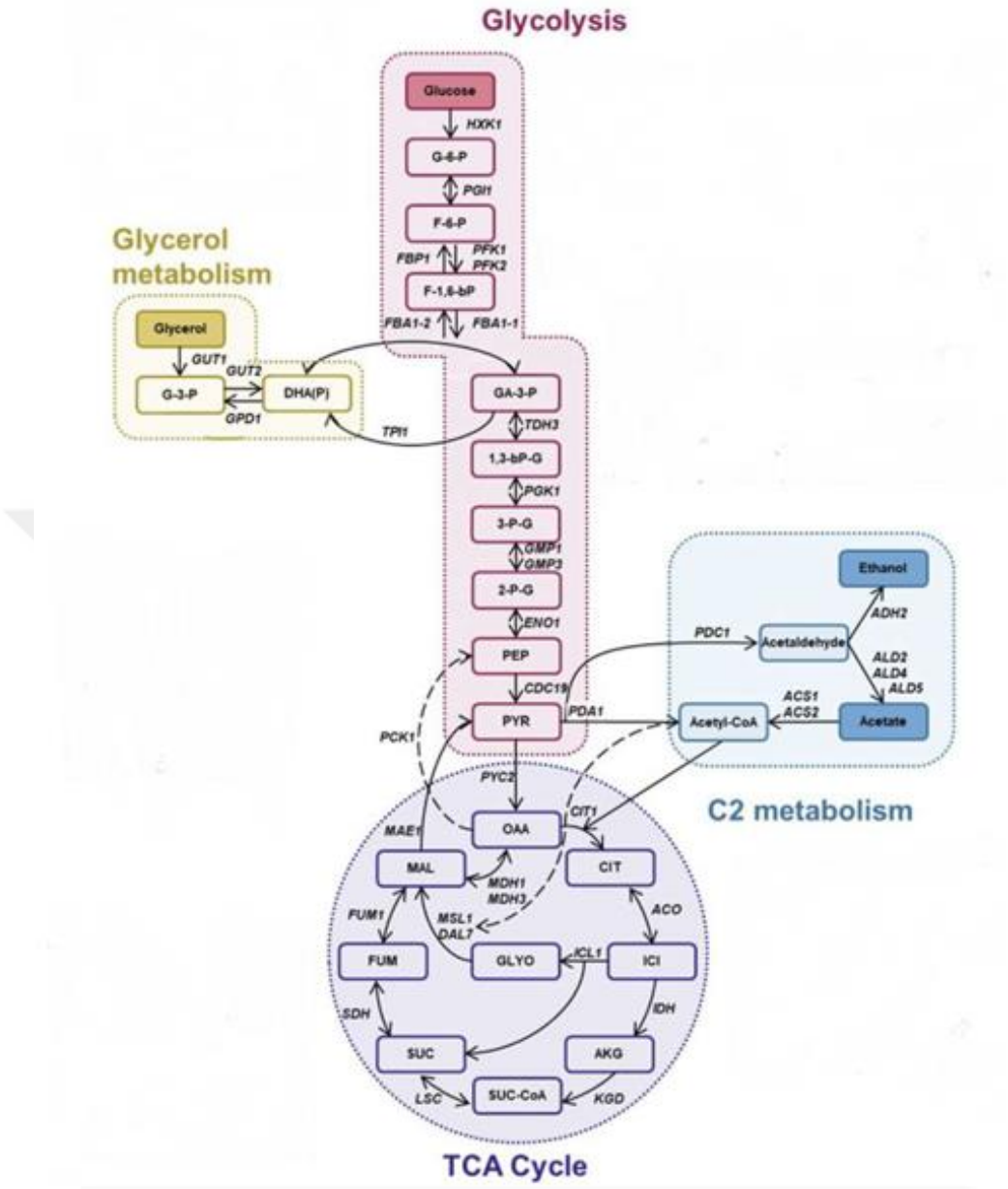


Figure 4. 43 Central carbon metabolism pathways of *P. pastoris* (Paes *et al.*, 2021)

The genes were selected because they are either crucial for the growth on gluconeogenic carbon sources involved in *P. pastoris* ethanol metabolism, or responsible for encoding regulatory transcriptional factors. Transcription levels of the selected genes for the wild-type strain, *ADH2*-OE-C3 and *ADH2*-OE-C9 were determined and presented in Figure 4.44. Thereby, both effects (i) *ADH2*

overexpression, and (ii) *ADH2* gene copy number, on the expression of genes involved in the ethanol utilization pathway, glyoxylate cycle, TCA cycle, and gluconeogenic genes were determined.

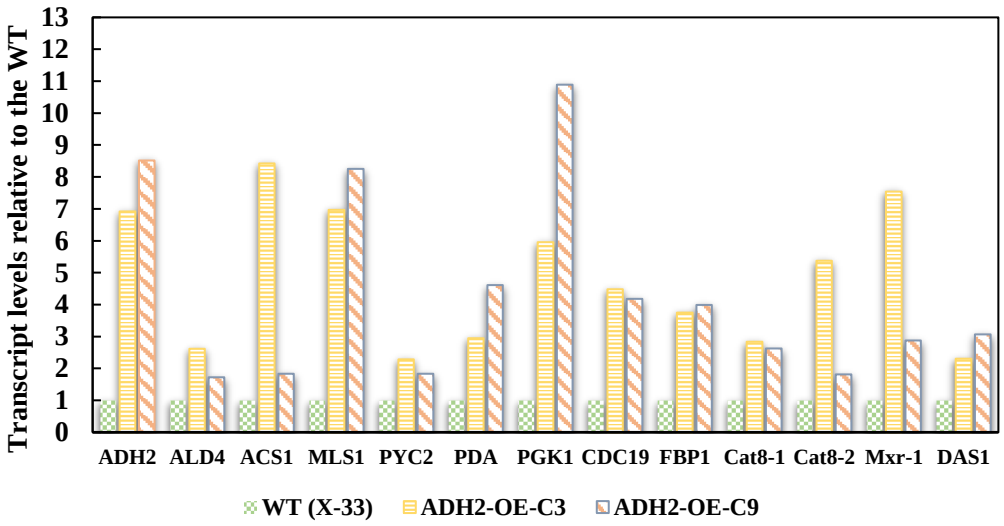


Figure 4. 44 Influence of *ADH2* overexpression on transcript levels of selected genes in *P. pastoris* induced on ethanol

The three genes encoding the crucial enzymes for assimilation of ethanol *ADH2*, *ALD4*, and *ACS1*, were upregulated in *ADH2* overexpression strains. A 6.9-fold increase was observed in the expression of *ADH2* in the colony containing three copies, while an 8.5-fold increase was obtained in the colony containing nine copies of *ADH2*. It is concluded that the increase in the copy number also increases the expression of *ADH2*, as expected. However, it should be noted that the increase in *ADH2* expression was not parallel with the gene copy number, which is in consistent with the fact that the decrease in expression with the increase in the gene copy number after a maximum should be a consequence of the pressure on the transcriptional machinery created by the tools used in the construction of the expression systems and their implementation. The increase in *ALD4* and *ACS1* transcription levels were inversely proportional to the *ADH2* copy number. In *ADH2*-C3 and *ADH2*-C9 strains, ALD4 was upregulated by 1.7-fold and 2.6-fold, respectively. This result is consistent with *ADH2* overexpression resulted in

accumulation of acetic acid and acetaldehyde. Since upregulation of *ADH2* is higher than both *ALD4* and *ACS1* in *ADH2*-C3 and *ADH2*-C9, accumulation of acetaldehyde and acetate occur. Acetate is sensed intracellularly, and as a result, ethanol metabolism causes the acetate response to be induced. Increasing *ADH2* gene copy number increased *ADH2* expression, resulting in higher acetaldehyde and acetic acid accumulation in *ADH2*-C9 strain. Inhibition of the cell growth is the outcome of toxic acetaldehyde and acetic acid accumulations. The expression levels of the *MLS1* gene, coding for the enzyme malate synthase, were analyzed to interpret the effects of *ADH2* overexpression on the glyoxylate cycle. *ADH2* overexpression upregulated the *MLS1* gene 6.9- and 8.2-fold in *ADH2*-C3 and *ADH2*-C9 strains, respectively. The glyoxylate cycle is indeed important for shuttling of acetyl-CoA after conversion of ethanol to acetyl-CoA over three consecutive reactions of the EUT pathway. The expression of *MLS1*, which participates in the utilization of non-fermentable carbon sources, is sensitive to carbon catabolite repression (Hartig et al., 1992). *MLS1* can accept butyryl-CoA as an acyl-CoA donor in addition to the traditional substrate acetyl-CoA. (Hartig et al., 1992; Kunze et al., 2002; Branduardi et al., 2013) Upregulation of *MLS1* showing that *ADH2* overexpression activated the glyoxylate cycle; consequently, the TCA cycle.

The malic enzyme catalyzes the oxidative decarboxylation of malate to pyruvate and carbon dioxide. Then, pyruvate is either converted to acetyl-CoA by the enzyme PDA1 or converted to oxalaacetate by the enzyme PYC2. High upregulation of *PDA1* indicates increased activation of C2 metabolism, while high expression of *PCY2* indicates activation of the TCA cycle. Both *ADH2*-C3 and *ADH2*-C9 strains upregulated the expression of *PDA1* and *PYC2*. Unlike *ADH2*-C3 strain, in *ADH2*-C9 strain, higher upregulation of *PDA1* and lower upregulation of *PYC2* was obtained. This implies that increasing *ADH2* gene copy number resulted in higher redirection of pyruvate to acetyl-CoA, which is converted either citrate with CIT1 or malate with MLS1. Since upregulation of *MLS1* is higher in *ADH2*-C9 than *ADH2*-C3, it is concluded that PYC2 is inhibited by malate accumulation in the cell. Furthermore, in *ADH2*-C3 strain and *ADH2*-C9 strain, CDC19, pyruvate kinase

enzyme which converts phosphoenolpyruvate to pyruvate, the input for aerobic TCA cycle (Mortimer et al., 1989), expression upregulated 4.5-fold and 4.1-fold, respectively. Since separate colonies with different copies of the *ADH2* gene resulted in the similar upregulation of *CDC19*, either a direct relationship between the gene copy number and expression of this gene has not been established or pyruvate is a rigid node so that the reaction rates from pyruvate to AcCoA is not affected. In fermentative and non-fermentative carbon source metabolism, gene expression patterns are very different. On ethanol, instead of the glycolysis pathway, the gluconeogenesis pathway should be active, furthermore either the TCA cycle is completed or glyoxylate cycle should be active or both. Thus, the expression levels of three gluconeogenic genes were investigated to understand the effect of *ADH2* overexpression on the upstream part of the central carbon metabolism. FBP1, Fructose-1,6-bisphosphatase; key regulatory enzyme in the gluconeogenesis pathway, expression upregulated in both *ADH2*-C3 and *ADH2*-C9 with similar increase as, 3.7 and 3.9. On the other hand, *PGK1* expression upregulated 5.6-fold and 10.9-fold in colonies *ADH2*-C3 strain and *ADH2*-C9 strain, respectively. PGK1, 3-phosphoglycerate kinase, catalyzes the transfer of high-energy phosphoryl groups from the acyl phosphate of 1,3-bisphosphoglycerate to ADP to produce ATP and is a key enzyme in gluconeogenesis. Consequently, it is concluded that there is a direct correlation between *ADH2* copy number and *PGK1* expression. Overall, activation of the gluconeogenesis is higher in the *ADH2* overexpression strain containing nine copies of ADH2 gene, compared with the three copies of ADH2-containing strain and the wild-type strain.

CAT8-1 and *CAT8-2* expression were also upregulated in *ADH2* overexpressed strains. This is in consistent with the fact that when Cat8 protein is activated via phosphorylation by Snf1, induces the transcription of genes involved in the growth on non-fermentable carbon sources such as the genes of the C2 anabolism, the glyoxylate cycle, and gluconeogenesis. (De Vit et al. 1997; Haurie et al. 2001; Hedges et al.1995; Lesage et al. 1996; Randez-Gil et al. 1997; Tachibana et al. 2005;

Kalender and Çalık 2020; Barbay et al., 2021,) Furthermore, *MXR1* expression was also upregulated in both colonies with respect to the wild-type strain. The pattern of increase in *MXR1* expression appears to be very similar to the pattern of upregulation in *ACS1*. Just as with the increase in *ACS1* expression, *MXR1* expression is higher in the *ADH2*-C9. An increase in *MXR1* expression is related to *ACS1* expression because Mxr1p regulates the expression of *ACS1* encoding acetyl-CoA synthetase 1 during acetate metabolism (Rao et al., 2020).

After analysis of *ADH2* overexpression and the number of copies of *ADH2* gene, the effects of *ACS1* overexpression on the expression of genes involved in ethanol pathway, glyoxylate cycle, TCA cycle, and gluconeogenic genes were investigated. Transcription levels of 13 genes are presented in Figure 4.45.

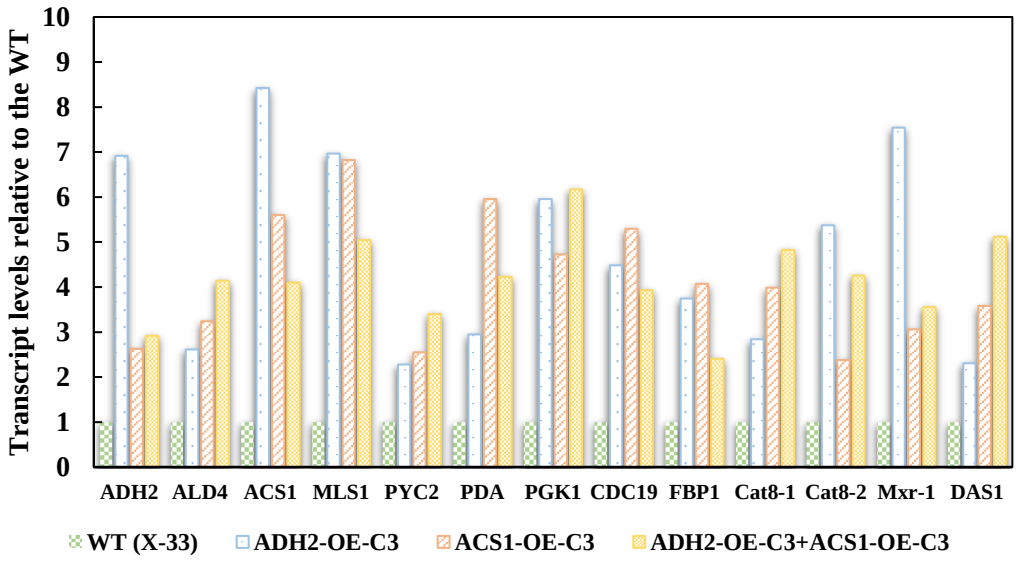


Figure 4. 45 Influence of *ADH2* and *ACS1* overexpression on transcript levels of selected genes in *P. pastoris* induced on ethanol

Three genes encoding the crucial enzymes in ethanol utilization, *ADH2*, *ALD4*, and *ACS1*, were upregulated in the *ACS1*-OE strain. A higher increase was obtained in *ADH2* gene expression in *ADH2*-OE than *ACS1*-OE, as expected. *ACS1* overexpression also resulted in the upregulation of all the three genes in the EUT

pathway. However, there appears to be an inverse correlation between *ALD4* upregulation and *ACS1* upregulation. The expression of the *ALD4* gene increased 2.61-fold and 3.23-fold in *ADH2*-OE and *ACS1*-OE, respectively. On the other hand, upregulation of *ACS1* was 8.42-fold and 5.59-fold in *ADH2*-OE and *ACS1*-OE. The gene expression ratios (ER) of enzymes of central metabolism obtained by *ADH2*-OE, *ACS1*-OE, *and ADH2*-OE +*ACS1*-OE were given in Figure 4.46.

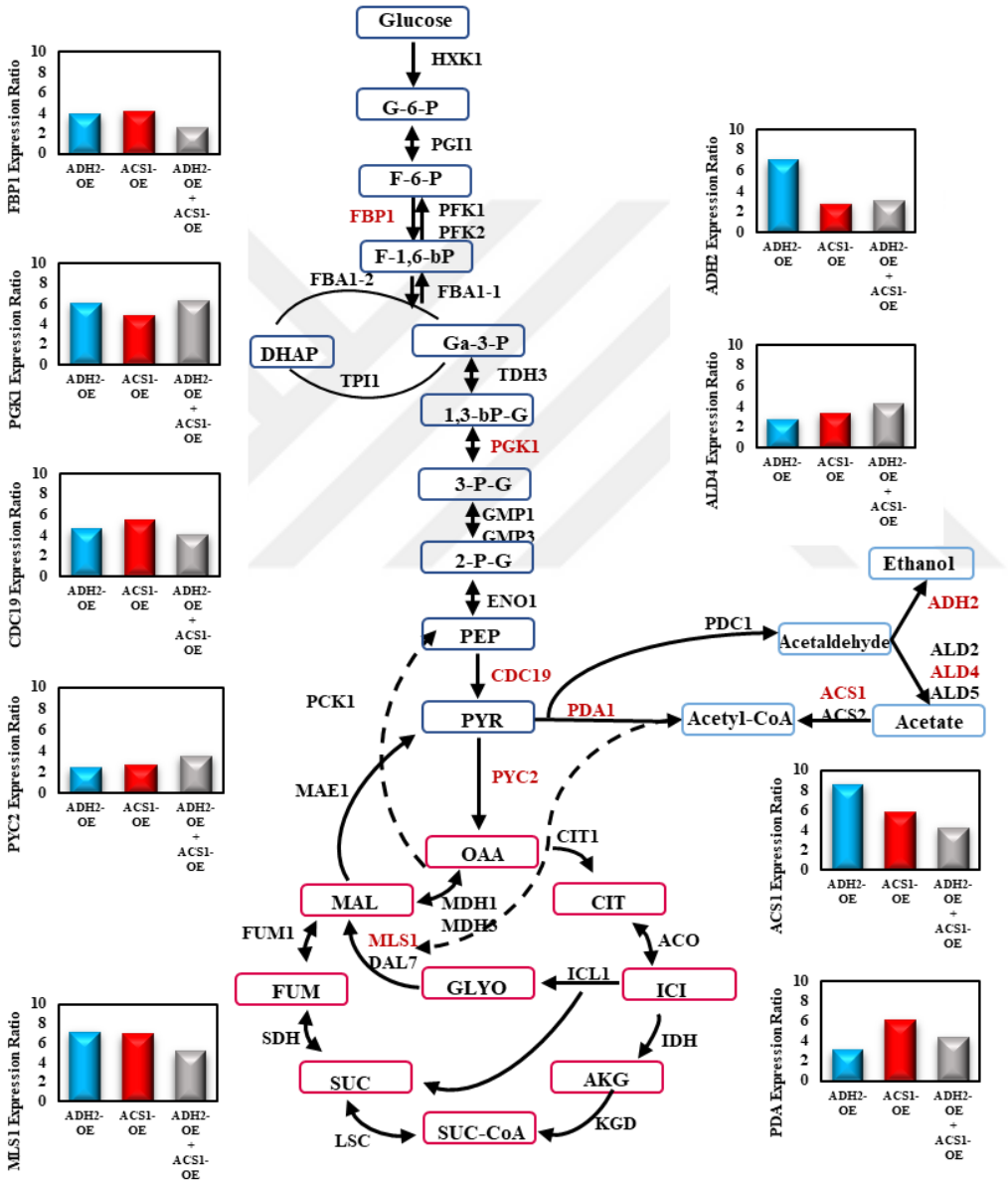


Figure 4. 46 The gene expression ratios (ER) based on housekeeping gene *ACT1* obtained by *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE + *ACS1*-OE in central metabolism.

MLS1 expression is higher for *ADH2* overexpression strain than *ACS1* overexpression strain. *MLS1* expression increased 7-fold for *ADH2*-OE and 6.8-fold for *ACS1*-OE. In contrast, the expression levels of the *PYC2* gene are inversely proportional to that of *MLS1*. For *ADH2* and *ACS1* overexpression strains, *PYC2* ratio increased 2.28 and 2.54-fold, respectively. This is in agreement with the result that is high malate concentrations in the cell cause decrease in the activation of *PYC2*. Pyruvate carboxylase is a cytoplasmic enzyme that converts pyruvate to oxaloacetate. In other words, *PYC2* redirected pyruvate produced from malate to the TCA cycle. Thus, higher transcription of *PYC2* implies higher fluxes in the TCA cycle. Furthermore, pyruvate is a branch point, so with the PDA1 enzyme pyruvate can be converted to acetyl-CoA, or by PDC enzyme pyruvate can be converted to acetaldehyde. Higher *PDA1* expression refers to higher acetyl-CoA synthesis, consequently upregulates the TCA cycle genes. *PDA1* expression upregulated in *ACS1*-OE strain 5.96-fold. In *ADH2* overexpression strain, *PDA1* was upregulated 2.95-fold. In the *ADH2*-OE + *ACS1*-OE strain, *PDA1* expression ratio was 4.22-fold. The increase in *PDA1* expression is related to the *ACS1* gene. For the *ACS1*-OE strain, 4.72-fold and 5.29-fold upregulation were obtained in the expression of PGK1 and CDC19 genes, respectively. Moreover, *PGK1* was upregulated 6-fold in the *ADH2*-OE strain.

Transcription factor concentrations are changed with the overexpression of *ADH2* and *ACS1,* since in the construction of overexpression cassettes with pADH2 and pAOX1/Cat8-L3 promoters were used, which contains Cat8 binding sites. Expressions of *CAT8-1* and *CAT8-2* were upregulated in both *ADH*-OE and *ACS1*-OE strains. To make a more general comparison including all overexpression systems generated, the expression levels of the selected genes were analyzed for three overexpression systems, including wild type and Venn diagrams of upregulation ratios of genes involved in the C2 metabolism, TCA cycle and gluconeogenesis, in *ADH2*-OE*, ACS1*-OE and *ADH2*-OE+*ACS1*-OE strains are presented in Figure 4.47, separately.

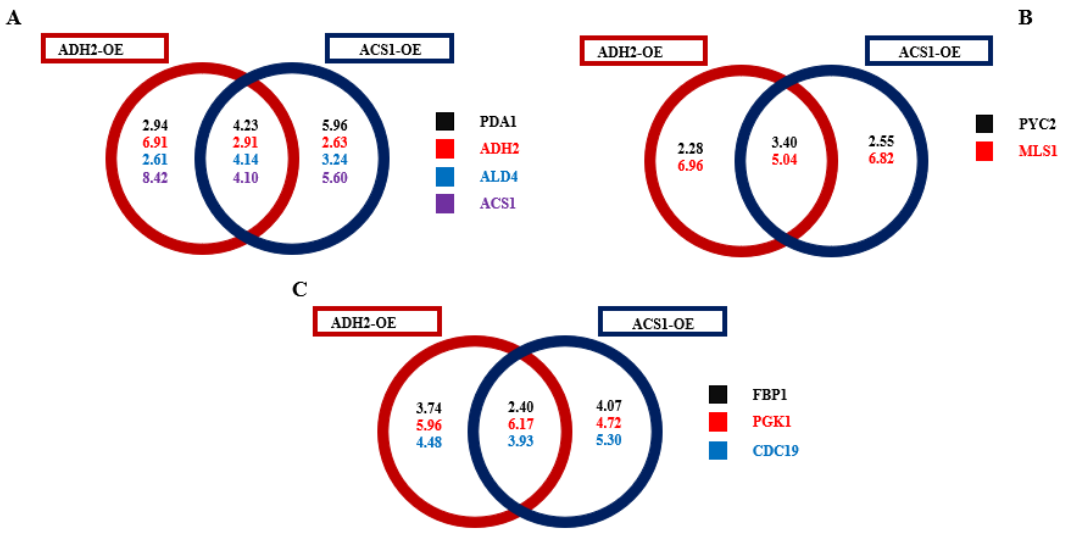


Figure 4. 47 Venn diagrams of upregulation ratios of genes in *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE+*ACS1*-OE strains. (A) Upregulation of genes involved in C2 metabolism; (B) Upregulation of genes involved in TCA cycle; (C) Upregulation of genes involved in gluconeogenesis.

Enzymes in the EUT pathway were upregulated in all the generated overexpression strains. The expression of the ALD4 gene increased 4.1-fold in the *ACS1*+*ADH2* overexpression strain, giving the highest upregulation. The highest expression of the *ACS1* gene was obtained with the *ADH2* overexpression strain. When the *ACS1*+*ADH2* overexpression strain results are considered, it is concluded that the expressions of *ALD* and *ACS1* genes increase at similar levels (4-fold). For the strain *ADH2*-OE + *ACS1*-OE, there is a 5-fold upregulation of *MLS1* expression. However, when the expression levels of the *PYC2* gene are examined, it is seen that it is inversely proportional to *MLS1*. For *ADH2* and *ACS1* overexpression strains, there was an increase of 2.28 and 2.54-folds, respectively, while this increase was 3.39-fold for *ACS1*+*ADH2* overexpression strains.

PDA1 expression upregulated in *ACS1*-OE strain 5.96-fold. On the other hand, in *ADH2* overexpression strain, 2.95-fold upregulation was obtained. Further, for the *ADH2*-OE + *ACS1*-OE strain, upregulation was 4.22-fold higher, indicating that

ACS1 overexpression increases *PDA1* expression more than *ADH2*. In other words, the increase in *PDA1* expression is interrelated with ACS1 gene expression.

On the other hand, the upregulation of the *CDC19* gene in *ADH2*-OE and *ADH2*-OE + *ACS1*-OE strains were 4.18-fold and 3.98-fold. Moreover, *PGK1* upregulated 6- and 6.16-fold in *ADH2*-OE and *ADH2*-OE + *ACS1*-OE strains. When the *PGK* and *CDC19* transcript levels of the three strains are analysed together, it is seen that the *PGK1* gene expression is higher in the strains containing *ADH2*-OE, and the *CDC19* is lower compared to *PGK1*. In the *ACS1* overexpression strains, *CDC19* upregulation is higher compared to *PGK1* upregulation. This means that the downstream activation of core carbon metabolism is higher for the strain *ACS1*-OE. In other words, high *PDA1* and *CDC19* expression for strain *ACS1*-OE indicate high regulation of the TCA cycle.

An interesting result obtained from the transcript data is that all of the controlled genes in all the strains, including the *DAS1* gene, which is one of the methanol metabolism genes, were upregulated compared to that of the X-33. The relevant explanation for this phenomenon is as follows: pAOX1/Cat8-L3 promoter was used to construct the *ACS1* overexpression cassette. Since promoter contains a Cat8 binding site, causes an upregulated transcription of *CAT8*, which increases the expression of all the genes regulated by Cat8.

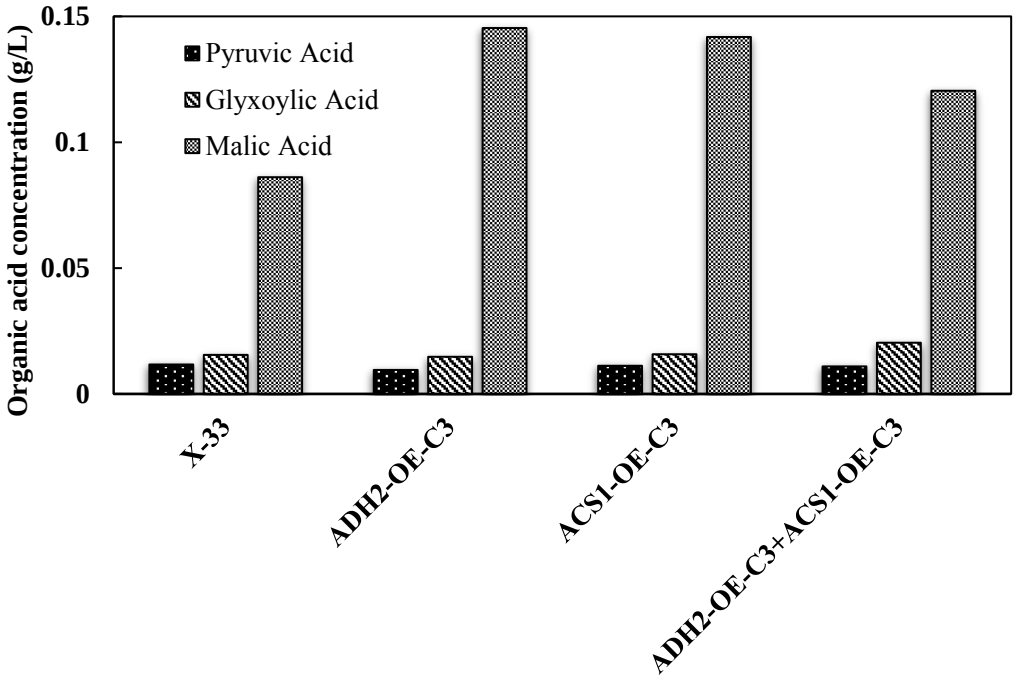


Figure 4. 48 Organic acid concentrations in the cell culture at t = 36 h for wild-type *P. pastoris X-33*, *ADH2*-OE, *ACS1*-OE, and *ADH2*-OE+*ACS1*-OE overexpression strains on 1% (v/v) ethanol

To uncover the relative contribution of enzyme and metabolite concentrations in controlling fluxes, metabolite concentrations and transcription levels of enzymes were analysed together. Compared to glyoxylic acid and pyruvic acid, higher malic acid concentrations were obtained with all generated overexpression strains (Figure 4.48). The highest and lowest malic acid concentrations were obtained in the *ADH2-OE* strain and *ADH2*-OE + *ACS1*-OE strains, respectively. As expected, *MLS1* expression and malic acid concentrations were consistent. An inverse relationship was identified between malic acid concentration and *PYC2* expression; the highest upregulation of *PYC2* was obtained in *ADH2*-OE + *ACS1*-OE, and also lowest malic acid concentration was obtained with *ADH2*-OE + *ACS1*-OE strain. This is in agreement with the fact that the accumulation of malic acid causes the inhibition of *PYC2*. Pyruvic acid concentrations were similar in all the generated strains, C_a. 0.01 g/L, indicating that demand and supply of pyruvic acid is almost equal. This also

implies that produced malate converted to oxaloacetate, and with the activation of *PCK1*, it is redirected to the gluconeogenesis pathway. Expressions of *PGK1*, an enzyme of lower the gluconeogenesis pathway, supports this result since the highest upregulation of *PGK1* was obtained in *ADH2*-OE as 5.96-fold. Finally, it was clarified that glyoxylic acid concentrations are consistent with an inverse relationship with *MLS1* expression levels. Since the highest *MLS1* upregulation and the lowest glyoxylic acid concentration were obtained with *ADH2*-OE, and the lowest *MLS1* upregulation and the highest glycolic acid concentration were obtained with *ADH2*-OE + *ACS1*-OE strain, respectively.

4.4 Effects of *ACS1* overexpression on r-protein production on ethanol

In order to investigate the effect of *ACS1* overexpression on r-protein production, a novel host strain was generated in which the gene of the reporter r-protein *mApple* was expressed under the control of $P_{ADH2\text{-}Cat8\text{-}L2}$, and *ACS1* was overexpressed under $P_{AOX1/Cat8\text{-}L3}$.

The constructed *ACS1* overexpression cassette was integrated into the recombinant *P. pastoris* cells containing a single copy of $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* that was generated in our research group (Demir and Çalık 2020). *ACS1* overexpression strains were constructed by the transformation of the monodirectional *ACS1* gene carrying vector into the recombinant *P. pastoris* cells expressing *mApple* under the control of EPV $P_{ADH2\text{-}Cat8\text{-}L2}$. $P_{AOX1/Cat8\text{-}L3}$::*ACS1* overexpression cassettes were integrated into recombinant host cells carrying one copy of $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* expression cassettes.

Experiments were conducted with $P_{AOX1/Cat8\text{-}L3}$::*ACS1*+ $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* strain, carrying three *ACS1* gene copies designated with *ACS1*-C3+mApple-C1, carrying four *ACS1* gene copies designated with *ACS1*-C4+mApple-C1 and five *ACS1* gene copies designated with *ACS1*-C5+mApple-C1. The strength of *ACS1* overexpression

colonies was analyzed and compared with wild-type *P. pastoris* X-33 and $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* strains. The expression strength of each colony was determined in terms of *mApple* expressions in batch cultivations at t = 24 h in ASMv6 medium on 0.5% (*v/v*) and 1% *(v/v)* ethanol. Reporter protein expressions on different concentrations of ethanol are given in Figure 4.49 comparatively with each recombinant host cell.

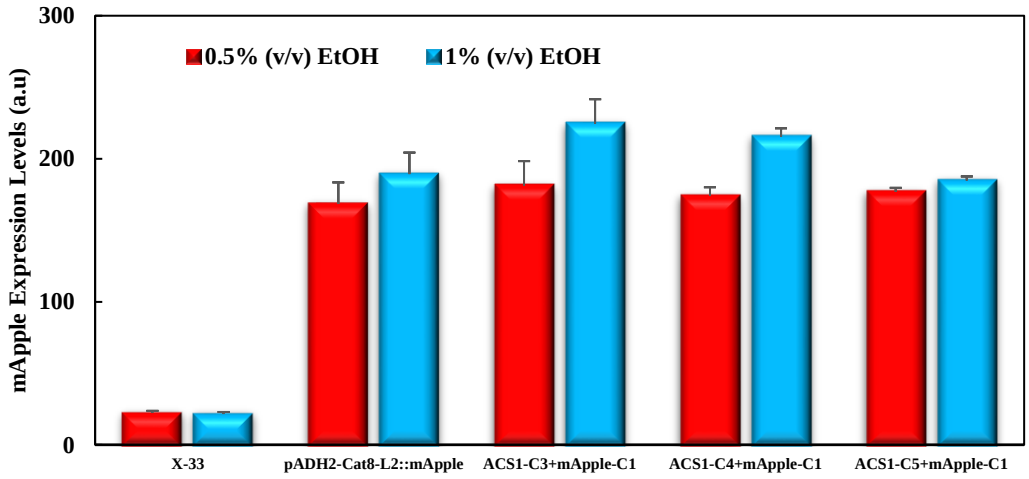


Figure 4. 49 mApple expressions of *P. pastoris* strains constructed with $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* and *ACS1* overexpression cassettes at t = 24 h of batch fermentation in ASMv6 medium on the carbon source of ethanol in V_R = 2-mL cultures in 12-deep well plates at T = 30 °C, pH_o = 5.0, N = 200 min^{-1}. Reporter FP mApple expressed under the EPV *ADH2-Cat8-L2*. Expressions with recombinant host cells and overexpression strains at 0.5 % (*v/v*) and 1 % (*v/v*), of ethanol. Error bars represent the standard deviation (±).

The highest mApple expression level was obtained on 1 % (*v/v*) ethanol in *ACS1*-C3+mApple-C1, which is 1.20-fold higher than the expression level in $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple*. A 1.14-fold increase in mApple expression was determined in *ACS1*-C4+mApple-C1. However, no significant increase was obtained in *ACS1*-C5+mApple-C1. The same pattern was also obtained in 0.5% (v/v) ethanol; however,

compared to 1% (v/v) ethanol, the increase in mApple expression was lower, 1.08-fold for *ACS1*-C3+mApple-C1 and 1.05-fold for *ACS1*-C4+mApple-C1. When mApple expression levels were considered together with the copy numbers, it was revealed that there is an inverse correlation between mApple expression levels and *ACS1* gene copy number. Increasing *ACS1* gene copy number resulted in a decrease in mApple expression. This finding is in agreement with the fact that the decrease in the expression with the increase in the gene copy number after a maximum should be the consequence of the pressure on the transcriptional machinery created by the tools used in the construction of the expression systems and their implementation (Demir and Çalık, 2020).

The results obtained at two ethanol concentrations showed that the production was increased with the increase in initial ethanol concentration. To generate production profiles of *ACS1*-C3+mApple-C1 and $P_{ADH2\text{-}Cat8\text{-}L2}$:: *mApple* and to compare ethanol consumption throughout the process, the fermentation process started, and time-dependent production profiles were generated by analyzing mApple expression levels at certain hours; in addition, specific ethanol uptake rate profiles were determined (Figure 4.45 and 4.46). Production experiment carried out with pADH2/Cat8-L2*::mApple* strain on 1% (*v/v*) ethanol to be able to compare the production capacity of *ACS1* overexpression strains.

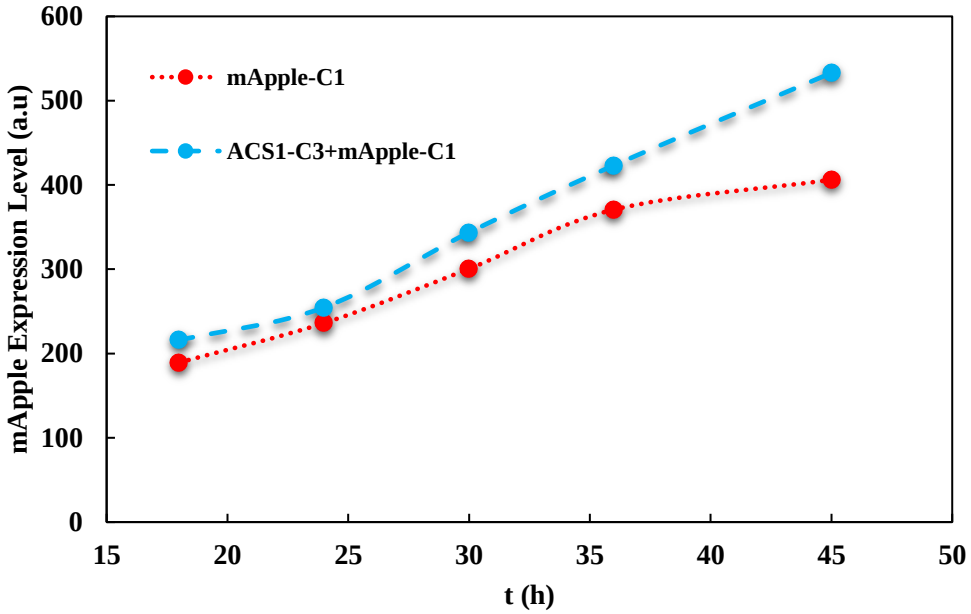


Figure 4. 50 Variations in the *mApple* expression levels with the cultivation time in air-filtered shake bioreactors on 1% (v/v) ethanol.

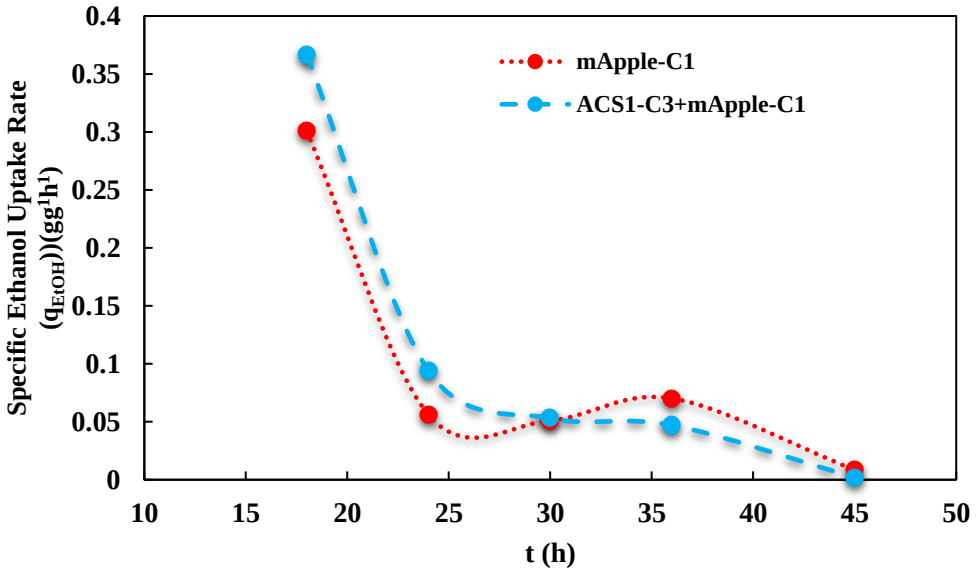


Figure 4. 51 Variations in specific ethanol uptake rates with the cultivation time in air-filtered shake bioreactors on 1% (v/v) ethanol.

The specific ethanol uptake rates (qEtOH) of *ACS1* overexpression strains for the first 18 h were higher than pADH2/Cat8-L2::mApple (Figure 4.46). At t= 18 h, mApple expression was also higher for the strains created with *ACS1* overexpression compared to the host cell. (Figure 4.45) Between 18 < t < 24 h, there was a significant decrease in the specific ethanol uptake rate for *ACS1*-C4+mApple-C1 and pADH2/Cat8-L2::*mApple* strains. Although the specific uptake rates decreased between 18 < t < 24 h, mApple expressions increased. The production and specific ethanol uptake rate profiles were similar in *ACS1*-C4+mApple-C1 and pADH2/Cat8-L2::*mApple*. Between 24 < t < 30 h, highest mApple expression was reached with *ACS1*-C4+mApple-C1. After t = 30 h, the specific ethanol uptake rate of pADH2/Cat8-L2::*mApple* reached higher values compared to *ACS1* overexpression strains. However, this situation was not reflected in mApple production performance, and although higher specific ethanol uptake rates were obtained for pADH2/Cat8-L2::mApple until the end of the process (t=45h), mApple expression levels were higher for *ACS1* overexpression strain. This can be interpreted as *ACS1*-OE colonies take ethanol into the cell in a shorter time and use it more effectively. With the faster uptake of ethanol, the reactions on the EUT pathway accelerated; thus, higher cell concentrations and mApple expression levels were achieved. At the end of the process, a 1.32-fold increase in mApple expression was obtained in *ACS1*-C4+mApple-C1.

4.5 *P. pastoris* strains with double-promoter expression system architectures

Double-promoter expression system (DPES) design as de novo metabolic engineering strategy enables fine-tuned and enhanced gene expression (Demir and Çalık, 2020). In this study, a collection of monodirectional hybrid-architectured DPESs were constructed with engineered promoter variant P_{mAOX1} and with the naturally occurring promoter P_{GAP} to enhance and upregulate-deregulated gene expressions in *Pichia pastoris* in methanol-free media. Reporter red fluorescent

protein (mApple) and enhanced green fluorescent protein (eGFP) were expressed under P_{mAOX1} or P_{GAP}, respectively, enabling the determination of the transcription period and strength of each constituent in the DPESs. We determined fluorescent protein expressions in batch cultivations on 1% (v/v) ethanol and excess glucose and compared them with single-promoter expression systems constructed with P_{mAOX1} and P_{GAP}.

4.5.1 Recombinant plasmids constructed with DPES and SPES in *P. pastoris*

With the aim of construction of DPES and SPES in *P. pastoris*, a recombinant plasmid containing the pmAOX::*mApple* cassette was constructed. For this purpose, pmAOX::*mApple* cassette was inserted into the vector-plasmid pADH2-Cat8-L2*::hGH*. For the selection based on NTC, pADH2-Cat8-L2*::hGH* plasmid, which was constructed by replacement of zeocin resistance gene with NTC resistance gene, was used as vector plasmid. To construct the DPES with two different expression cassettes by sequential transformation, a new antibiotic resistance gene other than Zeocine was required.

For the construction of plasmid, proper restriction enzymes, which are non-cutter of insert and one-site cutter of vector plasmid, were chosen as *Asc*I and *Xba*I. SnapGene and Restriction mapper was used to find the required restriction sites. In Figure 4.52, vector plasmid, pADH2-Cat8-L2*::hGH*, and restriction enzyme that were used in the construction of recombinant plasmid were given.

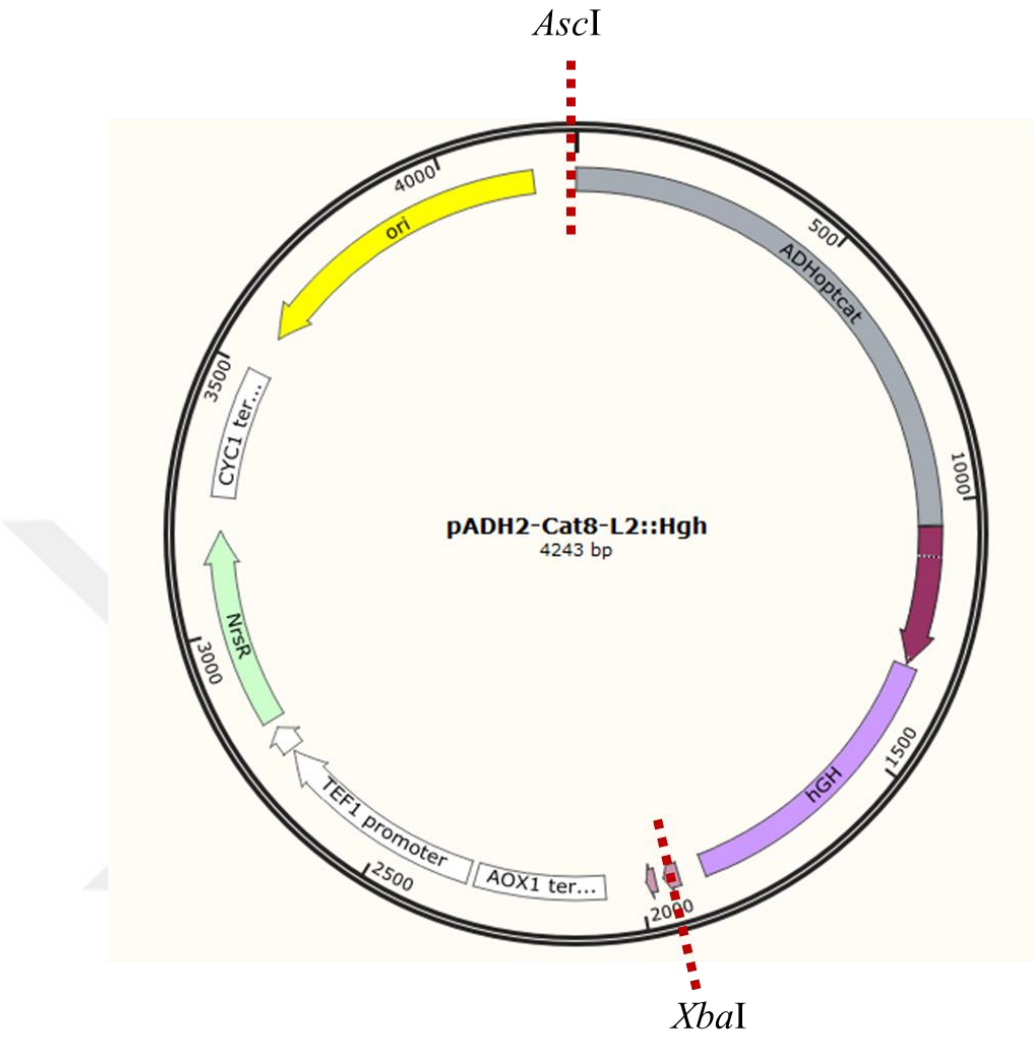


Figure 4. 52 Base plasmid, pADH2-Cat8-L2::*hGH*, and restriction enzymes, *Asc*I and *Xba*I, used in construction of recombinant plasmid.

The pmAOX1::*mApple* was constructed by a two-step splicing by overlap extension (SOE) PCR. In the first step of the two-step splicing by overlap extention (SOE) PCR, pmAOX and m*Apple*, pmAOX1 (940 bp) and *mApple* (717 bp) fragments were amplified by High-fidelity Q5 DNA Polymerase (NEB) with P_{mAOX}_Forward and P_{mAOX}_Forward and mApple_Forward and mApple_Reverse primers, respectively. For amplification of the pmAOX1 fragment, pmAOX1::*eGFP* plasmid (Ergun *et al.* 2020), was isolated from *E. coli* and used as a template, and for amplification of *mApple* fragment, pADH2-Cat8-L2::*mApple* plasmid (Demir and Çalık, 2020) was

isolated and used as a template. PCR reaction compositions and operation conditions are given in Tables 3.23, 3.24, and 3.25. The length of pmAOX1 and *mApple* were 940 bp and 717 bp, respectively. (See fig. 4.53)

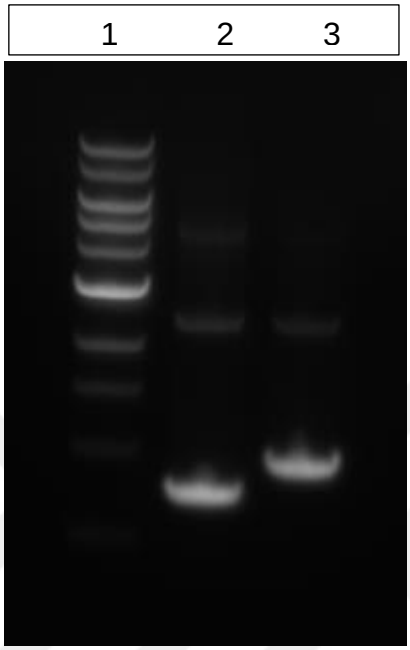


Figure 4. 53 Agarose gel electrophoresis of the gene sequences amplified with different primer combinations to construct pmAOX::*mApple* plasmid. 1: Quick-Load Purple 1 kb DNA ladder 2: P_{mAOX}_Forward and P_{mAOX}_Reverse; 3: *mApple*_Forward and mApple_Reverse.

The amplified fragments of P_{mAOX} and *mApple* was purified with GeneJet Gel Extraction Kit (ThermoFisher, USA) since other DNA fragments other than $P_{AOX\text{-}Cat8\text{-}L3}$ and *ACS1* fragments-were also amplified. To obtain purified $P_{AOX\text{-}Cat8\text{-}L3}$ and *A*CS1 fragments, gel extraction was conducted.

After purification of amplicons with overlapping compatible ends, they were used as template for the second step of SOE-PCR. Amplicon pairs purified in the first step of the SOE-PCR were coupled with primers P_{mAOX}_Forward and mApple_Reverse with the PCR mixture compositions and thermocycling conditions given in Table 3.26 and 3.27, respectively. $P_{AOX\text{-}Cat8\text{-}L3}$ and *ACS1* fragments fused for pAOX-Cat8-L3::*ACS1* synthesis and visualized in agarose gel electrophoresis. (Figure 4.54) The amplified pmAOX*::mApple* fragment is 1657 bp.

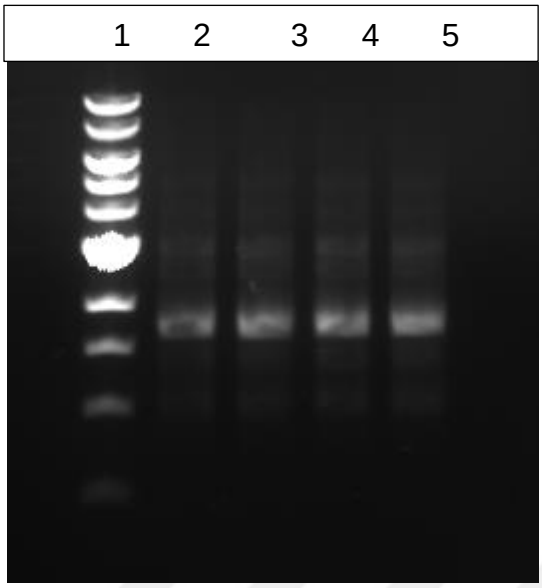


Figure 4. 54 Agarose gel electrophoresis of the gene sequences amplified with Forward_P_{mAOX} and Reverse_mApple to construct pmAOX::*mApple* fragment. 1: Quick-Load Purple 1 kb DNA ladder; 2-5: Amplified pmAOX::*mApple* fragment.

The amplified fragment of pmAOX::*mApple* was extracted from the gel by razor procedure, and to obtain purified pmAOX::*mApple* fragment gel extraction was conducted with GeneJet Gel Extraction Kit (ThermoFisher, USA). Purified fragment and vector plasmid, pADH2-Cat8-L2*::hGH,* were double digested with restriction enzymes *Asc*I and *Xba*I in a single restriction reaction and visualized in agarose gel electrophoresis (Figure 4.55)

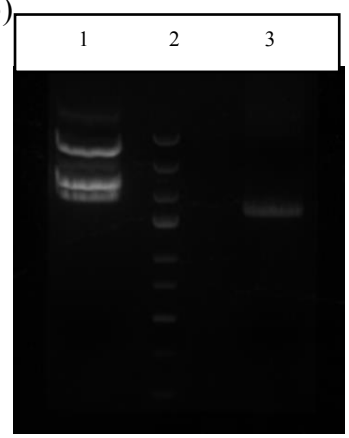


Figure 4. 55 Agarose gel electrophoresis of 1: Double digested vector with *Asc*I and *Xba*I; 2: GeneRuler Express; 3: Double digested insert, pmAOX::*mApple*, with *Asc*I and *Xba*I

The ligation reaction was performed to clone the double digested pAOX1-Cat8-L3::*ACS1* into double digested vector pGAPZαA using the vector: insert 3:1 ratio under the conditions specified in Section 3.4.1.4.

After completion of the ligation reaction, by the application of the calcium chloride method, transformation of the constructed plasmid into wild-type *E. coli* DH5α cells was performed. To make the selection, inoculation was done onto a YPD agar medium that contains kanamycin (Table 3.39). A single colony of *E. coli* DH5α transformed with pGAPZαA plasmids ligated with pADH2-wt::*ADH2* genes were used to inoculate 50 ml LB medium containing 50 μg/ml kanamycin in 250 ml.

After 16-18 hours of inoculation at 37°C, four colonies were selected from putative transformants and colony PCR, with Forward P_{mAOX}_Forward and mApple_Reverse, was carried out to control if the integration of plasmid was achieved, and visualized in agarose gel electrophoresis after colony PCR (Figure 4.56). The pmAOX1::*mApple* fragment is 1657 bp. Taq DNA Polymerase (Thermo Scientific) was used for the control PCR. The PCR operation conditions and reaction mixture contents of the colony PCR are given in Table 4.1 and 4.2, respectively.

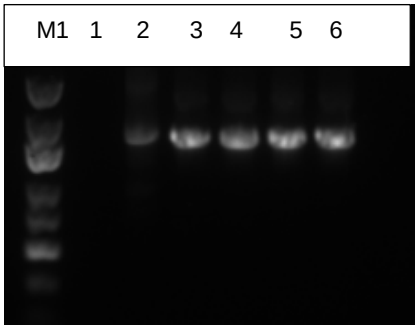


Figure 4. 56 Agarose gel electrophoresis of potential *E. coli* clones carrying the P_{mAOX}::*mApple* gene after transformation. M1: Quick-Load Purple 1 kb DNA

ladder; 1: Negative control for PCR control of insert pADH2-*ADH2* gene; 2: Positive control for PCR control of insert pADH2-*ADH2* gene, 3-6: PCR products of potential plasmids

As seen in Figure 4.56, all selected colonies are correct. Plasmid isolation was performed for two positive putative transformants to confirm the achievement of insertion of constructed plasmid, visualized in agarose gel electrophoresis after plasmid isolation (Figure 4.57).

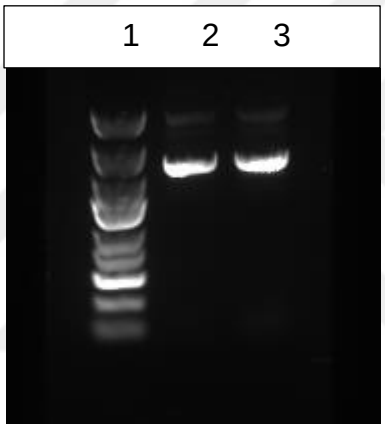


Figure 4. 57 Agarose gel electrophoresis image of potential recombinant plasmids. 1: Quick-Load Purple 1 kb DNA ladder, 2-5: Potential plasmids carrying the $P_{ADH2\text{-}wt}$::*ADH2* gene

After colony PCR, two colonies were selected, and constructs between *Bgl*II and *Kpn*I restriction sites were verified with DNA sequencing analysis (Sentebio, Ankara) (Appendix B). Glycerol stocks of *E. coli* strain that carry the verified constructed plasmid, pAOX1-Cat8-L3::*ACS1*, were prepared and stored at -80°C.

***P. pastoris* transformation**

Plasmids carrying the pmAOX::*mApple* cassette were linearized with Bam*HI* as specified in Table 3.33, and the linearized plasmid was extracted from the gel by the razor procedure purified with the gel elution kit, and visualized in agarose gel electrophoresis. (Figure 4.58)

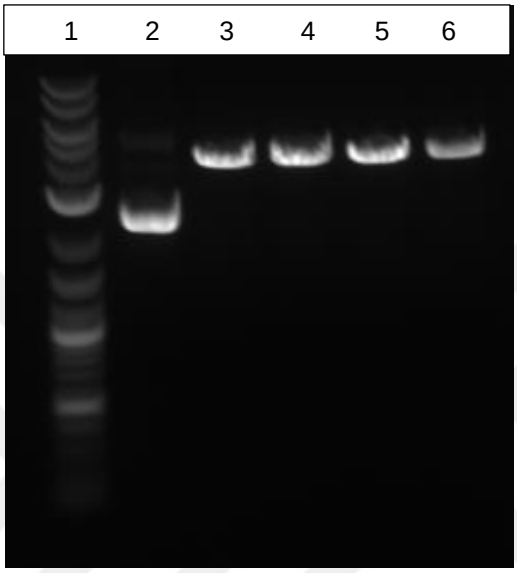


Figure 4. 58

Figure 4.58 Agarose gel electrophoresis image of linearized plasmid pmAOX1::*mApple* with *Bam*HI.. 1: Quick-Load Purple 1 kb DNA; 2: Circular pmAOX::*mApple* plasmid; 3-6: Linearized pmAOX::*mApple* plasmids.

The linearized plasmid was transformed into wild type X-33 *P. pastoris*, pADH2-Cat8-L2::*eGFP*, and pGAP::*eGFP* cells using the lithium chloride technique, as detailed in section 3.4.3. After transformation, 25-100 µl of cells were inoculated onto a selective YPD agar medium which contains 50 g/ml of NTC antibiotics and incubated for 48-60 hours at 30°C. Following the incubation, 12 colonies for each transformation were chosen from the growing cells, and the transfected gene was verified for integration using the PCR technique. According to colony PCR results, at least eight clones from each strain that were confirmed to carry the necessary gene were selected and inoculated into selective YPD media separately and incubated for

12-16 hours in air-filtered shake bioreactors at 30°C and 200 rpm. For using in future research, a glycerol stock was made from growing recombinant cells and preserved at -80°C.

4.5.2 Screening of Constructed DPESs and SPESs

After the transformation of *P. pastoris* wild-type X-33 and pADH2-Cat8-L2::*eGFP*, and pGAP::*eGFP* strains, at least eight clones were selected for each construct and analyzed in order to be able to select clones representing the entire population. To select the correct colonies representing the entire population, the screening experiments of the *P. pastoris* strains for intracellular eGFP, and mApple syntheses and extracellular rhGH production were carried out using 12-deep well plates. The randomly selected colonies were grown in 2 mL YPD medium containing 50 µg/mL NTC or 25 µg/ml Zeocin. *P. pastoris* clones were inoculated with an initial OD600 of 1.0 into the medium prepared with the carbon-source 1% (v/v) ethanol or 20 g/L excess glucose. The cells were harvested at t = 24 h of the cultivation and diluted with water to a final OD600 of 8.0 and were used for the measurement of mApple and eGFP expressions using a Fluorescence spectrophotometer (Agilent Technology, Cary Eclipse). Screening results of pmAOX*::mApple* are presented in Figure 4.59.

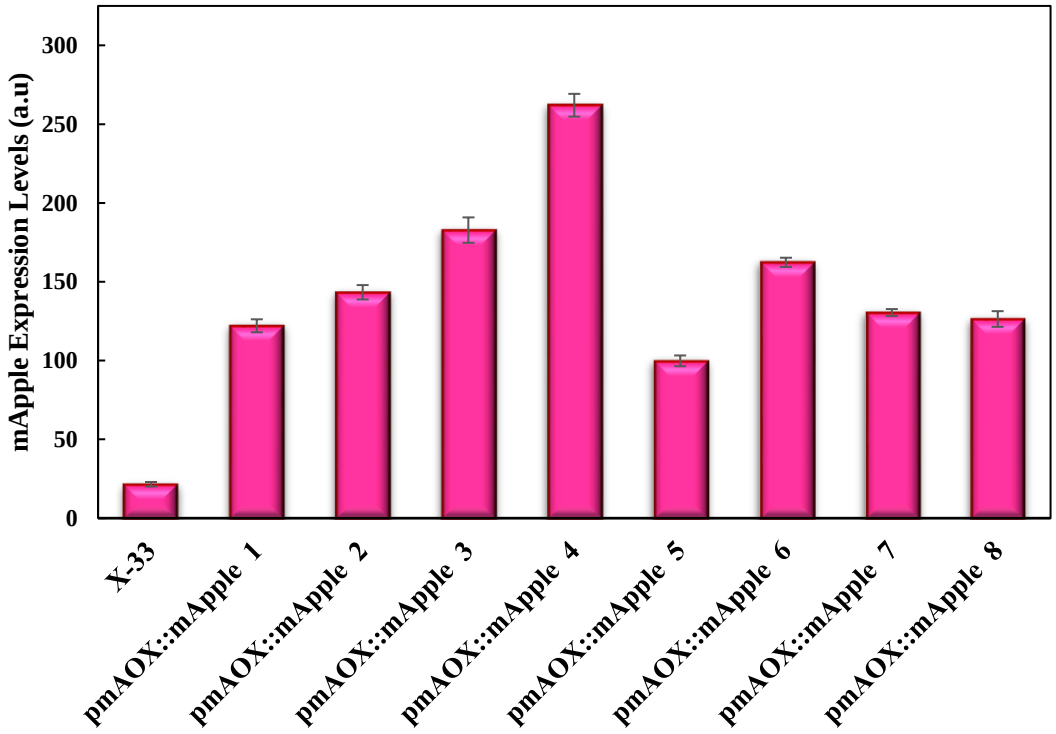


Figure 4. 59 mApple expression levels of *P. pastoris* strains carrying pmAOX::*mApple* on 1% (v/v) ethanol at t= 24 h of the fermentation. Error bars represent the standard deviation.

To construct the DPES with two different expression cassettes, the pmAOX::*mApple* plasmid was transformed into recombinant *P. pastoris* cells carrying pGAP::*eGFP* and. Six colonies and their single copies were screened on 1% (v/v) ethanol to find the true colony that represents the entire population. Two fluorescent proteins were used to identify the operation of the constituent NEPV (P_{mAOX1}) or NOP (P_{GAP}) and the operation period, separately, in the construction of the DPESs. As a result, when *mApple* was excited and emitted, P_{mAOX} expression is determined and quantified, or *eGFP* was excited and emitted, P_{GAP} expression is confirmed and quantified. pmAOX::mApple plasmid transformed to pGAP::*eGFP P. pastoris* cells by targeting the integration to the AOX1 locus. Expression performances of the biological replicates varies with the change in AOX1 locus which the constructed plasmids were integrated.

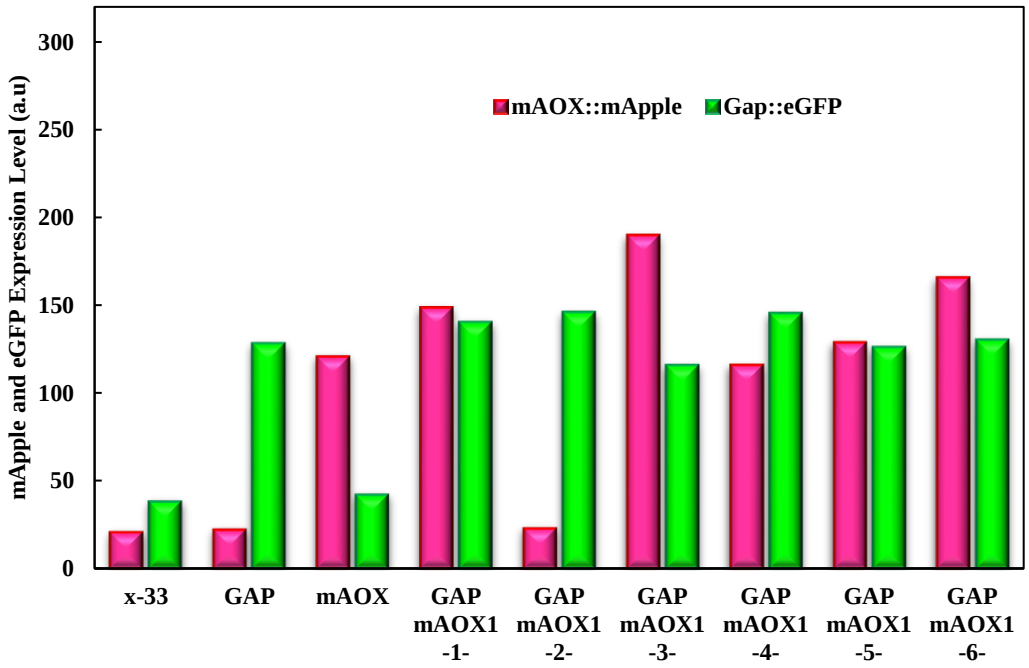


Figure 4. 60 *mApple* and *eGFP* expression levels of *P. pastoris* strains carrying the DPES pmAOX::*mApple* + pGAP::*eGFP* on 1% (v/v) ethanol at t=24 h of the fermentation. *mApple* and *eGFP* represent the expression levels of the constituent NEPV ADH2-Cat8-L2 and the NOP GAP, respectively.

Based on screening results of pADH2-Cat8-L2::*eGFP* + pmAOX::*mApple* DPES on 1% (v/v) ethanol, eGFP expression level with *GAP* promotor did not change among the colonies in the DPES as it is the parent-plasmid. Moreover, its expression capacity in the DPES was slightly increased compared with the SPES pGAP::*eGFP*. Similar mApple expression patterns were obtained with analyzed colonies except colony 3. The difference in the expressions in colony 3 probably resulted in due to increased gene-copy number. The differences in the expressions of the colonies compared to the SPES pmAOX::*mApple* might be caused by the variations in the integration site in the *AOX1* locus.

To generate production profiles of pmAOX::*mApple*+pGAP::*eGFP* DPES throughout the process, the fermentation process started, and time-dependent

production profiles were generated by analyzing mApple expression levels at certain hours.

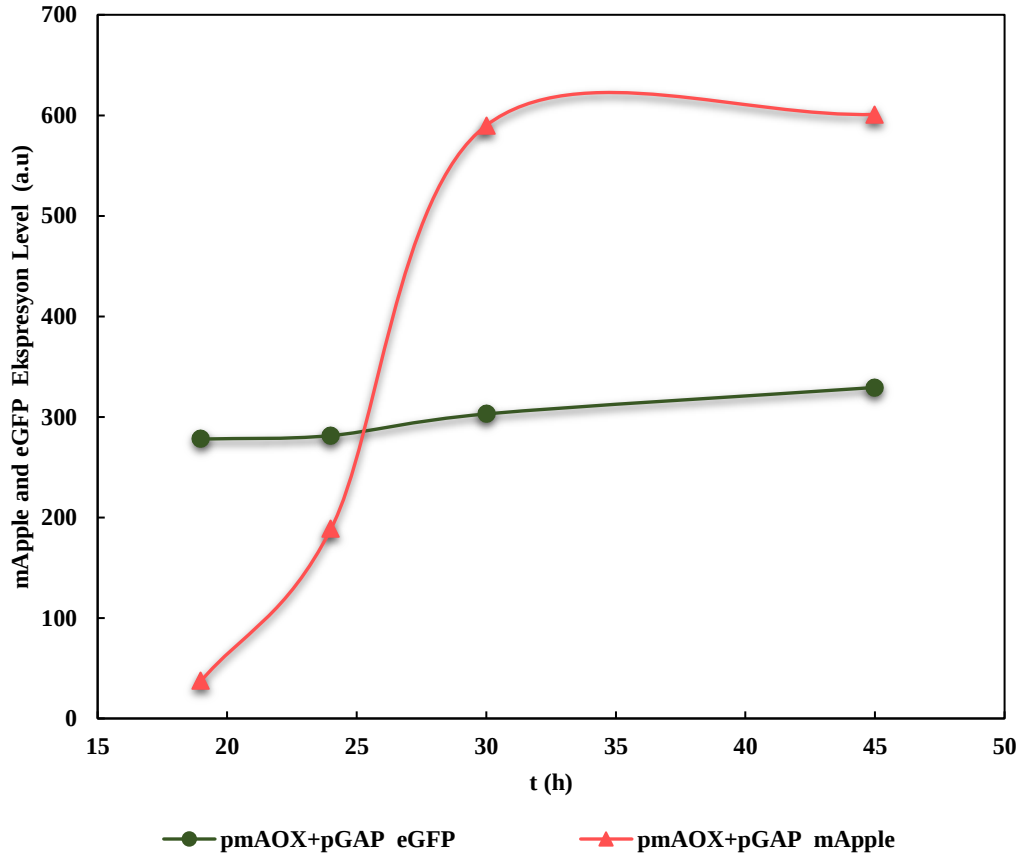


Figure 4. 61 Variations in the *mApple* and *eGFP* expression levels with the cultivation time in the air-filtered shake bioreactors on 20 g/L glucose.

The eGFP expression under the control of the P_{GAP} promoter in the glucose-containing production medium remained similar throughout the fermentation. A slight increase was observed at t= 30 h of the fermentation, and eGFP expression increased 1.08-fold. However, the expression level of mApple under the P_{mAOX1} of pmAOX+GAP DPES showed a significant increase from t=24 h of the process, which shows that the P_{mAOX} promoter was strongly induced by ethanol which is produced as a by-product in the fermentation process. At t=18 h, both glucose and ethanol concentration in the cell broth were analyzed to see whether ethanol was

produced after depletion of glucose. It was revealed that at t=18 h, glucose was completely depleted by the cells, and 0.3 g/L ethanol was produced. This result is in agreement with the fact that consumption of glucose resulted in ethanol production and induction of P_{mAOX1} promoter.

CHAPTER 5

CONCLUSIONS

To increase the ethanol uptake rate and decrease by-product formation in the ethanol utilization (EUT) pathway and acquire a smoothly operating intracellular reaction network, in the first part of this work, the alcohol dehydrogenase 2 (ADH2), and Acetyl-CoA synthetase (ACS1), two enzymes that catalyze the rate-limiting reactions, was overexpressed one by one and together and their effects on the ethanol uptake rate, cell growth, and by-product formation were investigated. To this end, novel metabolically engineered *Pichia pastoris* expression systems i) *ADH2*-OE, ii) *ACS1*-OE, iii) *ADH2*-OE + *ACS1*-OE were constructed. (Figure 5.1)

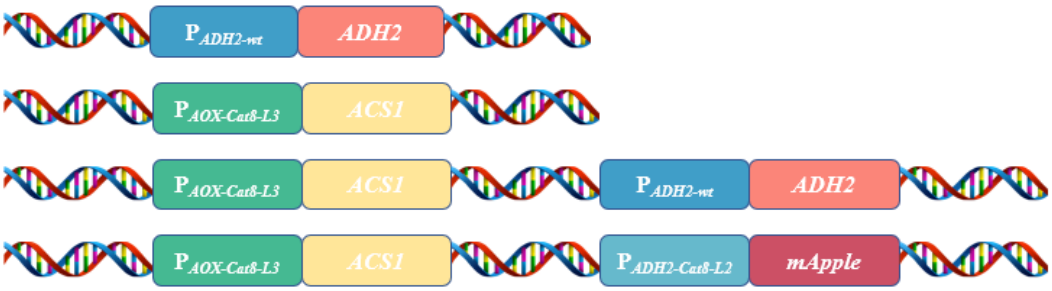


Figure 5. 1 Schematic representation of constructed *P. pastoris* expression systems.

With two separate colonies carrying 3 and 9 *ADH2* gene copies, the influence of *ADH2* overexpression on the growth of *P. pastoris* was investigated in batch cultivations on 2% (v/v) ethanol. The overexpression of *ADH2* resulted in growth inhibition; compared to *P. pastoris* X-33 1.1-fold and 1.35-fold lower cell concentrations were obtained with *ADH2*-OE-C3 and *ADH2*-OE-C9, respectively. Lower cell concentrations were associated with higher by-product formation in

metabolically engineered cells; at t = 36 h of the fermentation, compared to *P. pastoris* X-33, acetic acid concentrations were 1.46-fold and 3.97-fold higher in *ADH2*-OE-C3 and *ADH2*-OE-C9, respectively; while 1.18-fold and 2.28-fold higher acetaldehyde concentrations were obtained.

Metabolically engineered strains, i) *ADH2*-OE-C3, ii) *ACS1*-OE-C3, and iii) *ADH2*-OE-C3 + *ACS1*-OE-C3, were evaluated in terms of ethanol consumption, cell growth, and byproduct formation. On 1% (v/v) ethanol, with *ACS1*-OE compared to *ADH2*-OE and *ADH2-OE* + *ACS1-OE*, 1.08-fold and 1.12-fold higher cell concentrations were obtained, respectively; whereas, with *ACS1*-OE 1.78-fold lower acetic acid concentration was produced compared to *ADH2*-OE at t = 36 h of the fermentation. On the other hand, *ADH2* overexpression significantly increased the specific ethanol uptake rates. q_{EtOH} values at t = 9 h for *ADH2-OE* and *ADH2*-OE-C3 + *ACS1*-OE-C3 strains were 0.52 and 0.35 g g^{-1} h^{-1}, respectively. Compared to *P. pastoris* X-33, 0.93-fold lower cell concentrations were obtained-with *ACS1*-OE; furthermore, 1.49-fold and 1.07-fold lower acetic acid and acetaldehyde concentrations were obtained at t = 36 h of the fermentation, respectively.

In the second part of the study, the effects of modifications on the ethanol pathway on cell growth and ethanol uptake rate, the transcript levels of 13 genes in metabolically engineered strains on ethanol were analyzed, and changes in transcript levels in *P. pastoris* mutant strains were compared to wild-type strain using the threshold cycle (ΔΔCT) method. Three genes encoding crucial enzymes in ethanol utilization pathway, *ADH2*, *ALD4*, and *ACS1*, were upregulated in all three metabolically engineered strains. Moreover, *MLS1* expression was also upregulated in all constructed strains, and it was found that there is an inverse proportion between the upregulations of *MLS1* and *PYC2* genes. Higher malic acid concentrations compared to glyoxylic acid and pyruvic acid were obtained in all three metabolically engineered strains. Thus, it was concluded that higher malate concentrations in the cell cause a decrease in the activation of *PYC2*.

In the last part of the work, to investigate the effect of *ACS1* overexpression on r-protein production, a novel host strain was generated in which the gene of the reporter r-protein *mApple* was expressed under the control of $P_{ADH2\text{-}Cat8\text{-}L2}$, and *ACS1* was overexpressed under $P_{AOX1/Cat8\text{-}L3}$. At t = 24 h of the fermentation, *ACS1*-OE-C4 increased intracellular fluorescent protein synthesis 1.2-fold, and at t = 45 h of the fermentation, a 1.32-fold increase in florescent protein synthesis was obtained with *ACS1*-OE-C4.

REFERENCES

Aranda, A. and del Olmo, M. (2003), Response to acetaldehyde stress in the yeast *Saccharomyces cerevisiae* involves a strain-dependent regulation of several *ALD* genes and is mediatedby the general stress response pathway. Yeast, 20: 747-759.

Asada, H., Uemura, T., Yurugi-Kobayashi, T., Shiroishi, M., Shimamura, T., Tsujimoto, H., Ito, K., Sugawara, T., Nakane, T., Nomura, N., Murata, T., Haga, T., Iwata, S., & Kobayashi, T. (2011). Evaluation of the Pichia pastoris expression system for the production of GPCRs for structural analysis. *Microbial Cell Factories*, *10*, [24].

Baeshen, N. A., Baeshen, M. N., Sheikh, A., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A., ... & Redwan, E. M. (2014). Cell factories for insulin production. *Microbial cell factories*, *13*(1), 1-9.

Beier, D.R. and Young, E.T. (1982) Characterization of a regulatory region upstream of the *ADR2* locus of *S. cerevisiae*. *Nature*, **300** (5894), 724–728.

Berlec A, Strukelj B: Current state and recent advances in biopharmaceutical production in Escherichia coli, yeasts and mammalian cells. J Ind Microbiol Biotechnol 2013 http:// dx.doi.org/10.1007/s10295-013-1235-0.

Boubekeur, S., Camougrand, N., Bunoust, O., Rigoulet, M. & Guerin, B. Participation of acetaldehyde dehydrogenases in ethanol and pyruvate metabolism of the yeast Saccharomyces cerevisiae. *European journal of biochemistry* **268**, 5057–5065 (2001).

Çalik, P., Ata, Ö., Güneş, H., Massahi, A., Boy, E., Keskin, A., … Özdamar, T. H. (2015). Recombinant protein production in *Pichia pastoris* under glyceraldehyde-3-phosphate dehydrogenase promoter: From carbon source metabolism to bioreactor operation parameters. *Biochemical Engineering Journal*, *95*, 20–36. https://doi.org/10.1016/j.bej.2014.12.003

Carter PJ. Introduction to current and future protein therapeutics: a protein engineering perspective. Exp Cell Res 2011;317(9):1261–9.

Casatta, N., Porro, A., Orlandi, I., Brambilla, L. & Vai, M. Lack of Sir2 increases acetate consumption and decreases extracellular pro-aging factors. *Biochimica et biophysica acta* **1833**, 593–601, https://doi.org/10.1016/j.bbamcr.2012.11.008 (2013).

Çelik, E., & Çalik, P. (2012). Production of recombinant proteins by yeast cells. *Biotechnology Advances*, *30*(5), 1108–1118. https://doi.org/10.1016/j.biotechadv.2011.09.011

Cereghino, G. P. L., Cereghino, J. L., Ilgen, C., & Cregg, J. M. (2002). 1-s2.0-S0958166902003300-main. *Current Opinion in Biotechnology*, *13*, 329–332. https://doi.org/10.1016/S0958166902003300

Chen, Y., Daviet, L., Schalk, M., Siewers, V., & Nielsen, J. (2013). Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metabolic Engineering*, *15*, 48–54. doi: 10.1016/j.ymben.2012.11.002

Cheng, H., Lv, J., Wang, H. *et al.* Genetically engineered *Pichia pastoris* yeast for conversion of glucose to xylitol by a single-fermentation process. *Appl Microbiol Biotechnol* **98,** 3539–3552 (2014). https://doi.org/10.1007/s00253-013-5501-x

Choi, K. R., Jang, W. D., Yang, D., Cho, J. S., Park, D., & Lee, S. Y. (2019). Systems metabolic Engineering Strategies: Integrating systems and synthetic biology with metabolic engineering. *Trends in Biotechnology*, *37*(8), 817–837. https://doi.org/10.1016/j.tibtech.2019.01.003

Chudakov DM, Matz MV, Lukyanov S, Lukyanov KA. Fluorescent proteins and their applications in imaging living cells and tissues. Physiol Rev. 2010 Jul;90(3):1103-63. doi: 10.1152/physrev.00038.2009. PMID: 20664080.

Conner, J., Wuchterl, D., Lopez, M., Minshall, B., Prusti, R., Boclair, D., … Allen, C. (2014). *The Biomanufacturing of Biotechnology Products*. *Biotechnology Entrepreneurship: Starting, Managing, and Leading Biotech Companies*. Elsevier. https://doi.org/10.1016/B978-0-12-404730-3.00026-9

Corchero, J. L., Gasser, B., Resina, D., Smith, W., Parrilli, E., Vázquez, F., … Villaverde, A. (2013). Unconventional microbial systems for the cost-efficient production of high-quality protein therapeutics. *Biotechnology Advances*, *31*(2), 140–153. https://doi.org/10.1016/j.biotechadv.2012.09.001

Cormack, B. P., Valdivia, R. H., & Falkow, S. (1996). FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*, *173*(1), 33–38. https://doi.org/10.1016/0378-1119(95)00685-0

Corynebacterium glutamicum as analyzed by overexpression of the individual corresponding genes. Appl Environ Microbiol. 57 (1991) 1746−1752.

Costantini, L., Baloban, M., Markwardt, M. *et al.* A palette of fluorescent proteins optimized for diverse cellular environments. *Nat Commun* 6, 7670 (2015). https://doi.org/10.1038/ncomms8670

Couderc, R. and Baratti, J. (1980) Oxidation of methanol by the yeast Pichia pastoris: puri¢cation and properties of alcohol oxidase. Agric. Biol. Chem. 44, 2279^2289.

Cregg, J. (1993). The Pichia System. *Keck Graduate Institute, Claremont, Calif*, 1–8.

Cregg, J. M. (1999). 6 - Expression in the Methylotrophic Yeast Pichia pastoris. *Gene Expression Systems*, 157–191.

Cregg, J. M., Tolstorukov, I., Kusari, A., Sunga, J., Madden, K., & Chappell, T. (2009). *Chapter 13 Expression in the Yeast Pichia pastoris*. *Methods in Enzymology* (1st ed., Vol. 463). Elsevier Inc. https://doi.org/10.1016/S0076-6879(09)63013-5

de Smidt, O., du Preez, J. C., & Albertyn, J. (2008). The alcohol dehydrogenases of Saccharomyces cerevisiae: a comprehensive review. *FEMS yeast research*, *8*(7), 967–978. https://doi.org/10.1111/j.1567-1364.2008.00387.x

Demain AL, Vaishnav P: Production of recombinant proteins by microbes and higher organisms. Biotechnol Adv 2009, 27:297-306.

Demir, İ., & Çalık, P. (2020). Hybrid-architectured DOUBLE-PROMOTER expression systems enhance AND UPREGULATE-DEREGULATED gene expressions IN *Pichia Pastoris* in METHANOL-FREE MEDIA. *Applied Microbiology and Biotechnology*, *104*(19), 8381–8397. https://doi.org/10.1007/s00253-020-10796-5

Dueber, J. E., Wu, G. C., Malmirchegini, G. R., Moon, T. S., Petzold, C. J., Ullal, A. V., Prather, K. L., & Keasling, J. D. (2009). Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology*, *27*(8), 753–759. https://doi.org/10.1038/nbt.1557

Ergün, B. G., Demir, İ., Özdamar, T. H., Gasser, B., Mattanovich, D., & Çalık, P. (2020). Engineered deregulation of expression in yeast with designed hybrid-promoter architectures in coordination with discovered master regulator transcription factor. *Advanced Biosystems*, *4*(4), 1900172. https://doi.org/10.1002/adbi.201900172

Ergün, B. G., Gasser, B., Mattanovich, D., & Çalık, P. (2019). Engineering of alcohol DEHYDROGENASE 2 hybrid-promoter architectures IN *Pichia*

Pastoris to ENHANCE recombinant protein expression on ethanol. *Biotechnology and Bioengineering*, *116*(10), 2674–2686. https://doi.org/10.1002/bit.27095

Evdokimov AG, Pokross ME, Egorov NS, Zaraisky AG, Yampolsky IV, Merzlyak EM, Shkoporov AN, Sander I, Lukyanov KA, Chudakov DM. Structural basis for the fast maturation of Arthropoda green fluorescent protein. *EMBO Rep* 7: 1006–1012, 2006.

Fan, Y., Jimenez Del Val, I., Müller, C., Wagtberg Sen, J., Rasmussen, S. K., Kontoravdi, C., … Andersen, M. R. (2015). Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnology and Bioengineering*, *112*(3), 521–535. https://doi.org/10.1002/bit.25450

Faye, G., Leung, D.W., Tatchell, K., Hall, B.D., and Smith, M. (1981) Deletion mapping of sequences essential for in vivo transcription of the iso-1-cytochrome c gene. *Proc. Natl. Acad. Sci. U.S.A.*, **78** (4), 2258–2262.

Fernández F.J., López-Estepa M., Querol-García J., Vega M.C. (2016) Production of Protein Complexes in Non-methylotrophic and Methylotrophic Yeasts. In: Vega M. (eds) Advanced Technologies for Protein Complex Production and Characterization. Advances in Experimental Medicine and Biology, vol 896. Springer, Cham

Ferrer-Miralles, N., Domingo-espín, J., Corchero, J. L., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals, *8*, 1–8. https://doi.org/10.1186/1475-2859-8-17

Gellissen, Gerd et al. 1992. "Heterologous Protein Production in Yeast." *Antonie van Leeuwenhoek*.

Giannattasio, S., Guaragnella, N., Zdralevic, M. & Marra, E. Molecular mechanisms of Saccharomyces cerevisiae stress adaptation and programmed cell death in response to acetic acid. *Front Microbiol* **4**, 33, https://doi.org/10.3389/fmicb.2013.00033 (2013).

Giniger, E., Varnum, S.M., and Ptashne, M. (1985) Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell*, **40** (4), 767–774.

Guarente, L. (1984) Yeast promoters: positive and negative elements. *Cell*, **36** (4), 799–800.

Guarente, L. and Mason, T. (1983) Heme regulates transcription of the *CYC1* gene of *S. cerevisiae* via an upstream activation site. *Cell*, **32** (4), 1279–1286.

Guarente, L., Yocum, R.R., and Gifford, P. (1982) A *GAL10-CYC1* hybrid yeast promoter identifies the *GAL4* regulatory region as an upstream site. *Proc. Natl. Acad. Sci. U.S.A.*, **79** (23), 7410–7414.

Guilliermond A. 1920. Zygosaccharomyces pastori, nouvelle espèce de levures copulation hétérogamique. Bulletin de la Société Mycologique de France 36:203-11

Gustavsson, R. (2018). *Development of soft sensors for monitoring and control of bioprocesses*. Linköping: Linköping University Electronic Press.

H. Van Urk, W.S. Voll, W.A. Scheffers, J.P. Van Dijken, Transient-state analysis of metabolic fluxes in Crabtree-positive and Crabtree negative yeasts, Appl. Environ. Microbiol. 56 (1990) 281–287

Hackett, S. R., Zanotelli, V. R., Xu, W., Goya, J., Park, J. O., Perlman, D. H., Gibney, P. A., Botstein, D., Storey, J. D., & Rabinowitz, J. D. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. *Science*, *354*(6311). https://doi.org/10.1126/science.aaf2786

Hagman, A., Säll, T., Compagno, C., & Piskur, J. (2013). Yeast "Make-Accumulate-Consume" Life Strategy Evolved as a Multi-Step Process That Predates the Whole Genome Duplication. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0068734

Hahn, S., Hoar, E.T., and Guarente, L. (1985) Each of three" TATA elements" specifies a subset of the transcription initiation sites at the *CYC-1* promoter of *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U.S.A.*, **82** (24), 8562–8566.

Hahn, S., Young, E.T., and Hinnebusch, A. (2011) Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics*, **189** (3), 705–736.

Hartner, F. S., Ruth, C., Langenegger, D., Johnson, S. N., Hyka, P., Lin-Cereghino, G. P., Lin-Cereghino, J., Kovar, K., Cregg, J. M., & Glieder, A. (2008). Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic acids research*, *36*(12), e76. https://doi.org/10.1093/nar/gkn369

Heim, R., & Tsien, R. Y. (1996). Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology*, *6*(2), 178–182. https://doi.org/10.1016/s0960-9822(02)00450-5

Heiss S, Puxbaum V, Gruber C, Altmann F, Mattanovich D, Gasser B. 2015. Multistep processing of the secretion leader of the extracellular protein Epx1 in *Pichia pastoris* and implications for protein localization. Microbiology 161:1356–1368.

Heit, C., Martin, S. J., Yang, F. & Inglis, D. L. Osmoadaptation of wine yeast (Saccharomyces cerevisiae) during Icewine fermentation leads to high levels of acetic acid. *J Appl Microbiol* **124**, 1506–1520, https://doi.org/10.1111/jam.13733 (2018).

Imura, M., Nitta, K., Iwakiri, R., Matsuda, F., Shimizu, H., & Fukusaki, E. (2020). Comparison of metabolic pro fi les of yeasts based on the difference of the Crabtree positive and negative. *Journal of Bioscience and Bioengineering*, *129*(1), 52–58. https://doi.org/10.1016/j.jbiosc.2019.07.007

Inan, M., & Meagher, M. M. (2001). The effect of ethanol and ACETATE on protein expression IN *Pichia pastoris*. *Journal of Bioscience and Bioengineering*, *92*(4), 337–341. https://doi.org/10.1016/s1389-1723(01)80236-x

Ito, Y., Hirasawa, T., & Shimizu, H. (2014). Metabolic engineering of saccharomyces cerevisiae to improve succinic acid production based on metabolic profiling. *Bioscience, Biotechnology, and Biochemistry*, *78*(1), 151–159. https://doi.org/10.1080/09168451.2014.877816

J. Cremer, L. Eggeling, and H. Sahm, Control of the lysine biosynthesis sequence in
Jenkins N: Modifications of therapeutic proteins: challenges and prospects. Cytotechnology 2007, 53:121-125.

Kalender, Ö., & Çalık, P. (2020). Transcriptional regulatory proteins in CENTRAL CARBON metabolism OF *Pichia pastoris* and Saccharomyces cerevisiae. *Applied Microbiology and Biotechnology*, *104*(17), 7273–7311. https://doi.org/10.1007/s00253-020-10680-2

Karaoğlan, M, Erden-Karaoğlan, F, Yılmaz, S, İnan, M. Identification of major *ADH* genes in ethanol metabolism of *Pichia pastoris*. *Yeast*. 2020; 37: 227– 236. https://doi.org/10.1002/yea.3443

Karaoglan, M., Karaoglan, F. E., & Inan, M. (2016). Functional analysis of alcohol dehydrogenase (ADH) genes in *Pichia pastoris. BIOTECHNOLOGY LETTERS, 38(3), 463–469.*

Kogure T, Karasawa S, Araki T, Saito K, Kinjo M, MiyawakiA. A fluorescent variant of a protein from the stony coral *Montipora* facilitates dual-color single-laser fluorescence cross-correlation spectroscopy. *Nat Biotechnol* 24: 577–581, 2006.

Kratzer, S. & Schuller, H. J. Carbon source-dependent regulation of the acetyl-coenzyme A synthetase-encoding gene ACS1 from Saccharomyces cerevisiae. *Gene* **161**, 75–79 (1995).

Kulkarni R, 2016. Metabolic Engineering: Biological Art of Producing Useful Chemicals. Resonance, 21 (3), 233-237.

Kurtzman C. 2005. Description of Komagataella phaffii sp. nov. and the transfer of Pichia pseudopastoris to the methylotrophic yeast genus Komagataella. Int J Syst Evol Microbiol 55:973-6.

Liang, H., Ma, X., Ning, W., Liu, Y., Sinskey, A. J., Stephanopoulos, G., & Zhou, K. (2021). Constructing an ethanol utilization pathway in Escherichia coli to produce acetyl-CoA derived compounds. *Metabolic Engineering*, *65*, 223-231.

Liu, Y., Bai, C., Liu, Q., Xu, Q., Qian, Z., Peng, Q., Yu, J., Xu, M., Zhou, X., Zhang, Y., & Cai, M. (2019). Engineered ethanol-driven Biosynthetic system for improving production of acetyl-CoA DERIVED drugs in CRABTREE-NEGATIVE YEAST. *Metabolic Engineering*, *54*, 275–284. https://doi.org/10.1016/j.ymben.2019.05.001

M. Ikeda. Amino acid production processes, Adv Biochem Eng/Biotechnol. 79 (2003) 1−35.

M. Zimmer, Chem. Rev., 2002, 102, 759–781

Massahi, A., & Çalık, P. (2018). Naturally occurring novel promoters around pyruvate branch-point for recombinant protein production in *Pichia pastoris* (Komagataella phaffii): Pyruvate decarboxylase- and pyruvate kinase-promoters. *Biochemical Engineering Journal, 138*, 111-120.

Mattanovich D, Branduardi P, Dato L, Gasser B, Sauer M, Porro D: Recombinant protein production in yeasts. Methods Mol Biol 2012, 824:329-358

N. Milne, P. Thomsen, N. Mølgaard Knudsen, P. Rubaszka, M. Kristensen, L. Borodina (2020-07-01). "Metabolic engineering of Saccharomyces cerevisiae for the de novo production of psilocybin and related tryptamine derivatives"*. Metabolic Engineering. **60**: 25–36. doi:10.1016/j.ymben.2019.12.007. ISSN 1096-7176.*

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol* 20: 87– 90, 2002

Ogata, K., Nishikawa, H. and Ohsugi, M. (1969) A yeast capable of utilizing methanol. Agric. Biol. Chem. 33, 1519^1520.

Paes, B. G., Steindorff, A. S., Formighieri, E. F., Pereira, I. S., & Almeida, J. R. (2021). Physiological characterization and TRANSCRIPTOME analysis OF *Pichia PASTORIS* reveals its response To lignocellulose-derived inhibitors. *AMB Express*, *11*(1). https://doi.org/10.1186/s13568-020-01170-9

Pfeiffer, T., & Morley, A. (2014). An evolutionary perspective on the Crabtree effect, *1*(October), 1–6. https://doi.org/10.3389/fmolb.2014.00017

Phaff H, Miller M, Shifrine M. 1956. The taxonomy of yeasts isolated from Drosophila in the Yosemite region of California. Antonie Van Leeuwenhoek 22:145-61.

Piper, P., Calderon, C. O., Hatzixanthis, K., & Mollapour, M. (2001). Weak acid adaptation: the stress response that confers yeasts with resistance to organic acid food preservatives. *Microbiology (Reading, England)*, *147*(Pt 10), 2635–2642. https://doi.org/10.1099/00221287-147-10-2635

Prabhu, A.A., Veeranki, V.D. Metabolic engineering of *Pichia pastoris* GS115 for enhanced pentose phosphate pathway (PPP) flux toward recombinant human interferon gamma (hIFN-γ) production. *Mol Biol Rep* **45,** 961–972 (2018). https://doi.org/10.1007/s11033-018-4244-2

Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, Mattanovich D. 2013. Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. Microb. Cell Fact. 12:5.

Puxbaum V, Gasser B, Mattanovich D. 2016. The bud tip is the cellular hot spot of protein secretion in yeasts. Appl. Microbiol. Biotechnol. 100:8159–8168.

Qin X, Qian J, Yao G, Zhuang Y, Zhang S, Chu J (2011) GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. Appl Environ Microbiol 77(11):3600–3608. https://doi.org/10.1128/AEM.02843-10

R. Katsumata, T. Mizukami, Y. Kikuchi, K. Kino, Threonine production by the lysine producing strain of *Corynebacterium glutamicum*with amplified threonine biosynthetic operon. In: M. Alacevic, D. Hranueli, Z. Toman (eds) Genetics of Industrial Microorganisms.Yugoslavia (1986) p. 217.

Rhee, H.S. and Pugh, B.F. (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature*, **483** (7389), 295–301.

Rossanese OW, Soderholm J, Bevis BJ, Sears IB, Connor JO, Williamson EK, Glick BS. 1999. Golgi Structure Correlates with Transitional Endoplasmic Reticulum Organization in *Pichia pastoris* and Saccharomyces cerevisiae 145:69–81.

Ruth C, Zuellig T, Mellitzer a., Weis R, Looser V, Kovar K, Glieder a. 2010. Variable production windows for porcine trypsinogen employing synthetic inducible promoter variants in *Pichia pastoris*. Syst. Synth. Biol. 4:181–191.

S. Guillouet, A.A. Rodal, G. An, P.A. Lessard, A.J. Sinskey, Expression of the *Escherichia coli* catabolic threonine dehydratase in *Corynebacterium glutamicum* and its effect on isoleucine production. Appl Environ Microbiol. 65 (1999) 3100−3107

Sainsbury, S., Bernecky, C., and Cramer, P. (2015) Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.*, **16** (3), 129–143.

Saint-Prix, F., Bonquist, L. & Dequin, S. Functional analysis of the ALD gene family of Saccharomyces cerevisiae during anaerobic growth on glucose: the NADP+−dependent Ald6p and Ald5p isoforms play a major role in acetate formation. *Microbiology* **150**, 2209–2220, https://doi.org/10.1099/mic.0.26999-0 (2004).

Sakai Y, Koller A, Rangell LK, Keller GA, Subramani S. 1998. Peroxisome Degradation by Microautophagy in *Pichia pastoris*: Identification of Specific Steps and Morphological Intermediates J Cell Biol 141(3):625–636.

Sample, V., Newman, R. H., & Zhang, J. (2009). The structure and function of fluorescent proteins. *Chemical Society Reviews*, *38*(10), 2852. https://doi.org/10.1039/b913033k

Simpson-Lavy, K., Kupiec, M. Carbon Catabolite Repression in Yeast is Not Limited to Glucose. *Sci Rep* **9,** 6491 (2019). https://doi.org/10.1038/s41598-019-43032-w

Siripong, W., Wolf, P., Kusumoputri, T. P., Downes, J. J., Kocharin, K., Tanapongpipat, S., & Runguphan, W. (2018). Metabolic engineering of *pichia pastoris* for production of isobutanol and isobutyl acetate. *Biotechnology for Biofuels*, *11*(1). https://doi.org/10.1186/s13068-017-1003-x

Sjöblom M, Lindberg L, Holgersson J, Rova U. 2012. Secretion and expression dynamics of a GFP-tagged mucin-type fusion protein in high cell density *Pichia pastoris* bioreactor cultivations. Adv. Biosci. Biotechnol. 3:238–248..

Song, X., Li, Y., Wu, Y., Cai, M., Liu, Q., Gao, K., Zhang, X., Bai, Y., Xu, H., & Qiao, M. (2018). Metabolic engineering strategies for improvement of ethanol production in cellulolytic saccharomyces cerevisiae. *FEMS Yeast Research*, *18*(8). https://doi.org/10.1093/femsyr/foy090

Sreekrishna K., Kropp K.E. (1996) *Pichia pastoris*. In: Nonconventional Yeasts in Biotechnology. Springer, Berlin, Heidelberg

Stepanenko OV, Stepanenko OV, Shcherbakova DM, Kuznetsova IM, Turoverov KK, Verkhusha VV. Modern fluorescent proteins: from chromophore formation to novel intracellular applications. Biotechniques. 2011

Nov;51(5):313-4, 316, 318 passim. doi: 10.2144/000113765. PMID: 22054544; PMCID: PMC4437206.

Stepanenko, O. V., Verkhusha, V. V., Kuznetsova, I. M., Uversky, V. N., & Turoverov, K. K. (2008). Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes. *Current protein & peptide science*, *9*(4), 338–369. https://doi.org/10.2174/138920308785132668

Stephanopoulos, G. (2012). Synthetic biology and metabolic engineering. *ACS synthetic biology*, *1*(11), 514-525.

Stratford M., Anslow P. A. 1996; Comparison of the inhibitory action on *Saccharomyces cerevisiae* of weak-acid preservatives, uncouplers, and medium-chain fatty acids. *FEMS Microbiol Lett* 142:53–58

Stratford M., Anslow P. A. 1998; Evidence that sorbic acid does not inhibit yeast as a classic "weak acid preservative". *Lett Appl Microbiol* 27:203–206

Struhl, K. (1982) The yeast *his3* promoter contains at least two distinct elements. *Proc. Natl. Acad. Sci. U.S.A.*, **79** (23), 7385–7389.

Taymaz-Nikerel, H., Cankorur-Cetinkaya, A., and Kirdar, B. (2016). Genome-wide transcriptional response of *Saccharomyces cerevisiae* to stress-induced perturbations. *Front. Bioeng. Biotechnol.* 4:17. doi: 10.3389/fbioe.2016.00017

Tredwell, G. D., Aw, R., Edwards-Jones, B., Leak, D. J., & Bundy, J. G. (2017). Rapid screening of cellular stress responses in recombinant *pichia pastoris* strains using metabolite profiling. *Journal of Industrial Microbiology and Biotechnology*, *44*(3), 413–417. https://doi.org/10.1007/s10295-017-1904-5

Tripathi, N. K., & Shrivastava, A. (2019). Recent Developments in Bioprocessing of Recombinant Proteins : Expression Hosts and Process Development, *7*(December). https://doi.org/10.3389/fbioe.2019.00420

Tsiftsoglou, A. S., Ruiz, S., & Schneider, C. K. (2013). Development and regulation of biosimilars: Current status and future challenges. *BioDrugs*, *27*(3), 203–211. https://doi.org/10.1007/s40259-013-0020-y

use of its promoter. Gene 1997;186:37–44.

Van den Berg, M. A. & Steensma, H. Y. ACS2, a Saccharomyces cerevisiae gene encoding acetyl-coenzyme A synthetase, essential for growth on glucose. *European journal of biochemistry* **231**, 704–713 (1995).

Venters, B.J., Wachi, S., Mavrich, T.N., Andersen, B.E., Jena, P., Sinnamon, A.J., Jain, P., Rolleri, N.S., Jiang, C., Hemeryck-Walsh, C., and Pugh, B.F. (2011)

A comprehensive genomic binding map of gene and chromatin regulatory proteins in *Saccharomyces*. *Mol. Cell*, **41** (4), 480–492.

Vogl, T., Hartner, F. S., & Glieder, A. (2013). New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Current Opinion in Biotechnology*, *24*(6), 1094–1101. https://doi.org/10.1016/j.copbio.2013.02.024

Walsh G, Jefferis R: Post-translational modifications in the con- text of therapeutic proteins. Nat Biotechnol 2006, 24:1241-1252.

Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. Isolation of the Pichia pastoris glyceraldehyde-3-phosphate dehydrogenase gene and regulation and

Wehbe, O., Yaman, O. U., & Çalık, P. (2020). Ethanol fed-batch BIOREACTOR operation to enhance therapeutic protein production IN *Pichia pastoris* under hybrid-architectured ADH2 PROMOTER. *Biochemical Engineering Journal*, *164*, 107782. https://doi.org/10.1016/j.bej.2020.107782

Wen, J., Tian, L., Liu, Q., Zhang, Y., & Cai, M. (2020). Engineered dynamic distribution of malonyl-CoA flux for IMPROVING polyketide biosynthesis in Komagataella phaffii. *Journal of Biotechnology*, *320*, 80–85. https://doi.org/10.1016/j.jbiotec.2020.06.012

Wuest, D. M., Hou, S., & Lee, K. H. (2011). Metabolic Engineering. *Comprehensive Biotechnology*, 617–628. https://doi.org/10.1016/b978-0-08-088504-9.00229-4

Xiong L, Kang R, Ding R, Kang W, Zhang Y, Liu W, Huang Q, Meng J, Guo Z. Genome-wide Identification and Characterization of Enhancers Across 10 Human Tissues. Int J Biol Sci. 2018 Jul 27;14(10):1321-1332. PMID: 30123079; PMCID: PMC6097485.

Yamada Y, Matsuda M, Maeda K, Mikata K. 1995. The phylogenetic relationships of methanol-assimilating yeasts based on the partial sequences of 18S and 26S ribosomal RNAs: the proposal of Komagataella gen. nov. (Saccharomycetaceae). Biosci Biotechnol Biochem 59:439-44.

Yang, S.-T., Liu, X., & Zhang, Y. (2007). Metabolic Engineering – Applications, Methods, and Challenges. *Bioprocessing for Value-Added Products from Renewable Resources*, 73–118. https://doi.org/10.1016/b978-044452114-9/50005-0

Yang, Y.T., Bennet, G. N., San, K.Y., (1998) *Genetic and Metabolic Engineering*, Electronic Journal of Biotechnology, ISSN 0717-3458

Yu J, Xiao J, Ren X, Lao K, Xie XS. Probing gene expression in live cells, one protein molecule at a time. *Science* 311: 1600–1603, 2006.

Yu, X.-W., Sun, W.-H., Wang, Y.-Z., & Xu, Y. (2017). Identification of novel factors ENHANCING recombinant protein production IN multi-copy Komagataella PHAFFII based ON TRANSCRIPTOMIC analysis of Overexpression effects. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-16577-x

Yuan SF, Brooks SM, Nguyen AW, Lin WL, Johnston TG, Maynard JA, Nelson A, Alper HS. Bioproduced Proteins On Demand (Bio-POD) in hydrogels using *Pichia pastoris*. Bioact Mater. 2021 Jan 27;6(8):2390-2399. doi: 10.1016/j.bioactmat.2021.01.019. PMID: 33553823; PMCID: PMC7846901.

Zhang, Z. and Dietrich, F.S. (2005) Mapping of transcription start sites in *Saccharomyces cerevisiae* using 5' SAGE. *Nucleic Acids Res.*, **33** (9), 2838–2851. Pelechano, V., Wei, W., and Steinmetz, L.M. (2013) Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature*, **497** (7447), 127–131.

Zhu, T., Guo, M., Tang, Z., Zhang, M., Zhuang, Y., Chu, J., & Zhang, S. (2009). Efficient generation of multi-copy strains for optimizing secretory expression of porcine insulin precursor in yeast*pichia pastoris*. *Journal of Applied Microbiology*, *107*(3), 954–963. https://doi.org/10.1111/j.1365-2672.2009.04279.x

APPENDICES

A. Promoter and gene sequences

TEF promoter Hygromycin TEF Terminator

AGCTCAGGGGCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATC
CCCATGTATAATCATTTGCATCCATACATTTTGATGGCCGCACGGCGCG
AAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAAACGC
TCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGA
AATATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCC
TTTTAAAATCTTGCTAGGATACAGTTCTCACATCACATCCGAACATAAA
CAACCATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGT
TTCTGATCGAAAAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGA
GGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTGGATAT
GTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATG
TTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGA
CATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCA
CAGGGTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTC
TGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCTGCGGCCGATCTTA
GCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTCAAT
ACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTA
TCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCA
GGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGG
CACCTCGTGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATG
GCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGATGTTCGGGGATT
CCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTTG
TATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGC
AGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAA
CTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGC
AGGGTCGATGCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGC
GTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGATGGCTGTG
TAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGA
GGGCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAA
GAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATC
AAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAG
ATGCGAAGTTAAGTGC

> $P_{ADH2\text{-}wt}$

TCCTTTTTACCACCCAAGTGCGAGTGAAACACCCCATGGCTGCTCTCCG
ATTGCCCCTCTACAGGCATAAGGGTGTGACTTTGTGGGCTTGAATTTTA
CACCCCCTCCAACTTTTCTCGCATCAATTGATCCTGTTACCAATATTGCA
TGCCCGGAGGAGACTTGCCCCCTAATTTCGCGGCGTCGTCCCGGATCGC
AGGGTGAGACTGTAGAGACCCCACATAGTGACAATGATTATGTAAGAA
GAGGGGGGTGATTCGGCCGGCTATCGAACTCTAACAACTAGGGGGGTG
AACAATGCCCAGCAGTCCTCCCCACTCTTTGACAAATCAGTATCACCGA
TTAACACCCCAAATCTTATTCTCAACGGTCCCTCATCCTTGCACCCCTCT
TTGGACAAATGGCAGTTAGCATTGGTGCACTGACTGACTGCCCAACCTT
AAACCCAAATTTCTTAGAAGGGGCCCATCTAGTTAGCGAGGGGTGAAA
AATTCCTCCATCGGAGATGTATTGACCGTAAGTTGCTGCTTAAAAAAAA
TCAGTTCAGATAGCGAGACTTTTTTGATTTCGCAACGGGAGTGCCTGTT
CCATTCGATTGCAATTCTCACCCCTTCTGCCCAGTCCTGCCAATTGCCCA
TGAATCTGCTAATTTCGTTGATTCCCACCCCCCTTTCCAACTCCACAAAT
TGTCCAATCTCGTTTTCCATTTGGGAGAATCTGCATGTCGACTACATAA
AGCGACCGGTGTCCGAAAAGATCTGTGTAGTTTTCAACATTTTGTGCTC
CCCCCGCTGTTTGAAAACGGGGGTGAGCGCTCTCCGGGGTGCGAATTCG
TGCCCAATTCCTTTCACCCTGCCTATTGTAGACGTCAACCCGCATCTGGT
GCGAATATAGCGCACCCCCAATGATCACACCAACAATTGGTCCACCCCT
CCCCAATCTCTAATATTCACAATTCACCTCACTATAAATACCCCTGTCCT
GCTCCCAAATTCTTTTTTCCTTCTTCCATCAGCTACTAGCTTTTATCTTAT
TTACTTTACGAAA

ADH2

ATGTCTCCAACTATCCCAACTACACAAAAGGCTGTTATCTTCGAGACCA
ACGGCGGTCCCCTAGAGTACAAGGACATTCCAGTCCCAAAGCCAAAGT
CAAACGAACTTTTGATCAACGTTAAGTACTCCGGTGTCTGTCACACTGA
TTTGCACGCCTGGAAGGGTGACTGGCCATTGGACAACAAGCTTCCTTTG
GTTGGTGGTCACGAAGGTGCTGGTGTCGTTGTCGCTTACGGTGAGAACG
TCACTGGATGGGAGATCGGTGACTACGCTGGTATCAAATGGTTGAACG
GTTCTTGTTTGAACTGTGAGTACTGTATCCAAGGTGCTGAATCCAGTTG
TGCCAAGGCTGACCTGTCTGGTTTCACCCACGACGGATCTTTCCAGCAG
TATGCTACTGCTGATGCCACCCAAGCCGCCAGAATTCCAAAGGAGGCT
GACTTGGCTGAAGTTGCCCCAATTCTGTGTGCTGGTATCACCGTTTACA
AGGCTCTTAAGACCGCTGACTTGCGTATTGGCCAATGGGTTGCCATTTC
TGGTGCTGGTGGAGGACTGGGTTCTCTTGCCGTTCAATACGCCAAGGCT
CTGGGTTTGAGAGTTTTGGGTATTGATGGTGGTGCCGACAAGGGTGAAT
TTGTCAAGTCCTTGGGTGCTGAGGTCTTCGTCGACTTCACTAAGACTAA
GGACGTCGTTGCTGAAGTCCAAAAGCTCACCAACGGTGGTCCACACGG
TGTTATTAACGTCTCCGTTTCCCCACATGCTATCAACCAATCTGTCCAA
TACGTTAGAACTTTGGGTAAGGTTGTTTTGGTTGGTCTGCCATCTGGTGC
CGTTGTCAACTCTGACGTTTTCTGGCACGTTCTGAAGTCCATCGAGATC
AAGGGATCTTACGTTGGAAACAGAGAGGACAGTGCCGAGGCCATCGAC

TTGTTCACCAGAGGTTTGGTCAAGGCTCCTATCAAGATTATCGGTCTGT
CTGAACTTGCTAAGGTCTACGAACAGATGGAGGCTGGTGCCATCATCG
GTAGATACGTTGTGGACACTTCCAAATAA

P*AOX1/Cat8-L3*

AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCC
GACATCCACAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAG
GGGATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTC
TTCTCCTCAACACCCACTTTTGCCATCGAAAAACCAGCCCAGTTATTGG
GCTTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACAC
CATGACTTTATTAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTCATGTT
TGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGAACATCACTC
CAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAA
ATGGCCCAAAACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACA
AAAGCGTGATCTCATCCAAGATGAACTAAGTTTGGTTCGTTGAAATGCT
AACGGCCAGTTGGTCAAAAAGAAACTTCCAAAAGTCGGCATACCGTTT
GTCTTGTTTGGTATTGATTGACGAATGCTCAAAAATAATCTCATTAATG
CTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCGGTGCACCTGTGCCGA
AAC**ATATTCCGTTCGTCCGAAT**CTTTTTGGATGATTATGCATTGTCTCC
ACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAAC
GTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATA
TAAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATTA
TTAGCTTACTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATT
TTAACGACTTTTAACGACAACTTGAGAAGATCAAAAAACAACTAATTAT
TCGAAACG

ACS1

ATGCCATTAGATAACGAACACTTACTTCATGAAAATTCCATTGACCCAC
CAAAGGGATTCTTTGAAAGACACCCTGGAACTCCTAATATACCAGGCG
GTTGGGAAGAATACTTGAAGCTGTACAATCAGTCCATCGAGAACCCCTC
AAAGTTTTTTGGAGAAAAAGCAAAGGAATTCTTGTCATGGGCTACTCCT
TTCACTGACGCTCGTTACCCACCTGGTAATGGATTTCAGAATGGTGACT
CCGCCGCTTGGTTTCTGAATGGTGAGTTGAACGCGTCGTACAACTGTGT
TGATAGACATGCTTTAAAGAATCCAGACAAACCTGCCATTATTTATGAG
GCTGATGAACCTAATCAAGGCCGTACGGTTACCTATGGAGAGTTGCTGA
AGGATGTTTGTCGAATTGCCCAAGTATTGACTGACCTGGGTGTGAAAAA
GGGTGACACTGTTGCTGTTTACCTGCCTATGGTTCCAGAAGCTATCACC
ACTTTATTGGCTATCGTTAGAATCGGTGCTATCCACTCTGTTGTCTTCGC
AGGTTTTTCAGCTGGTTCTCTACGTGATCGTATATTGGATGCTGATTCTA
GAATTGTTATCACTTCTGATGAATCTCTGAGAGGTGGGAAGATCATCGA
GACTAAGAAGATTGTTGACGAGGCTCTGAAGTCTTGCCCAGATGTTCGT

AATGTGCTGGTCTTCAAAAGAACAGGTACACCACATCTTCCATGGGTTG
AGGGTCGTGATCTTTGGTGGCACGAGGAAATCATTAAGCATGTTCCGTA
CTCTCCCCCAGTGAATGTTAGATCTGAAGATACTTCATTTTTGCTTTACA
CTTCTGGCTCTACCGGAAAGCCTAAAGGTATCCAGCATTCAACTGCTGG
CTACTTACTGGGAGCTCTTTTGACCACCAAGTATGTCTTTGATGTTCAGG
GTGATGATATTTTATTCACTGCTGGTGATGTGGGCTGGATCACAGGGCA
TTCTTATGTAGTTTACGGTCCACTTTTAAACGGGGCTACGACAGTTGTTT
TTGAGGGCACCCCAGCTTACCCAGACTATTCACGTTATTGGGATATCGT
TGACAAACACAAAGTTACTCAGTTTTATGTAGCACCAACTGCTCTTAGG
TTGCTGAAGAGAGCTGGTAGCAAGTATGTCCAGAATCATGATTTGTCTT
CAATCAGGGTTTTGGGTTCCGTTGGTGAACCTATAGCCGCTGAAGTTTG
GGAATGGTACAACGAGTATGTTGGAAGAGGAAAAGCTCATATTTGTGA
TACGTATTGGCAAACAGAGACTGGTTCTCACATTATTGCTCCAATAGCT
GGTGTGTCAAAGACCAAACCAGGTTCAGCATCTTTCCCCTTCTTCGGTA
TTGATCCGGTTATTCTAGATGCTACTACTGGAGAGGAACTCAAAGGTAA
TAATGTTGAAGGTGTTTTGGCTATCAGAAATCCATGGCCATCTATGGCT
AGAACAGTCTGGAAGGACTACAACCGTTTCCTGGATACATATCTCAGGC
CATATGAAGGTTATTACTTCACTGGTGATGGAGCTGCCAGAGATCAGGA
AGGATTTTATTGGGTTCTGGGTAGAGTTGATGATGTTGTTAATGTGTCA
GGTCACAGATTGTCTACTGCCGAGATTGAAAGCGCTCTAATCGAACACA
ATTTGGTAGGAGAGTCTGCTGTCGTCGGATTCCCTGACGAGCTGACTGG
TTCTGCTGTGGCCGCGTTTGTGTCTTTGAAGAAGGACGTCGACAATCCA
GCGGAAGTGAAAAAGGAGTTAATCCTTACTGTCAGAAAAGAGATTGGA
CCATTCGCTGCACCTAAACTCATCATCTTGGTAAGTGATCTTCCAAAGA
CCAGATCAGGTAAGATAATGAGACGTATTCTCAGAAAGGTTTTGGCTG
GAGAGGAAGACTCTCTGGGCGACATTTCAACTCTTTCAAACCCTTCGAT
TGTGGAAGAGATAATCTCTACCGTTAAAAGGGATGCCCGCAAATGA

> P*mAOX1*

AGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGA

CATCCACAGGTCCATTCTCACACATAAGTGCCAAACGCAACA**ATTCCGTT**

CGTCCGATTAGCAGACCGTTGCAAACGCAGGACCTCCAC**ACCCCAATATT**

ATTTGGGGTACTTTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATT

GGAGCTCGCTCATTCCAATTCCTTCTATTAGGCTACTAACACCATGACTTTA

TTAGCCTGTCTATCCTGGCCCCCCTGGCGAGGTTCATGTTTGTTTATTTCCG

AATGC**CCTCTCGTCCGGGCTTTTT**CCGAACATCACTCCAGATGAGGGC**G**

ACCCCACATTTTTTTTTTGACCCCACATGTTCCCCAAATGGCCCAAAACT

GACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATCTCAT

CCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGT**GCCTATT**

GTAGACGTCAACCCAAGTCGGCATACCGTTTGTCTTGTTTGGTATTGATTG

ACGAATGCTCAAAAATAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGC
TTCTGAACCCCGGTGCACCTGTGCCGAAAC**ATATTCCGTTCGTCCGAAT**C
TTTTTGGATGATTA**ACCCCAATACATTTTGGGGT**TGCTTCCAAGATTCTGG
TGGGAATACTGCTGATAGCCTAACGTTCATGATCAAAATTTAACTGTTCTA
ACCCCTACTTGACAGCAATATATAAACAGAAGGAAGCTGCCCTGTCTTAAA
CCTTTTTTTTTATCATCATTATTAGCTTACTTTCATAATTGCGACTGGTTCCA
ATTGACAAGCTTTTGATTTTAACGACTTTTAACGACAACTTGAGAAGATCA
AAAAACAACTAATTATTCGAAACG

mApple

atggtgagcaagggcgaggagaataacatggccatcatcaaggagttcatgcgcttcaaggtgcacatggagggctc
cgtgaacggccacgagttcgagatcgagggcgagggcgagggccgcccctacgaggcctttcagaccgctaagct
gaaggtgaccaagggtggcccctgcccttcgcctgggacatcctgtcccctcagttcatgtacggctccaaggtctac
attaagcacccagccgacatccccgactacttcaagctgtccttccccgagggcttcaggtgggagcgcgtgatgaac
ttcgaggacggcggcattattcacgttaaccaggactcctccctgcaggacggcgtgttcatctacaaggtgaagctgc
gcggcaccaacttcccctccgacggccccgtaatgcagaagaagaccatgggctgggaggcctccgaggagcgga
tgtaccccgaggacggcgccctgaagagcgagatcaagaagaggctgaagctgaaggacggcggccactacgcc
gccgaggtcaagaccacctacaaggccaagaagcccgtgcagctgcccggcgcctacatcgtcgacatcaagttgg
acatcgtgtcccacaacgaggactacaccatcgtggaacagtacgaacgcgccgagggccgccactccaccggcg
gcatggacgagctgtacaagTAATCTAGA

Arg Standard Amplicon Size: 1259 bp

>XM_002490002.1 Komagataella phaffii GS115 Argininosuccinate lyase, catalyzes the final step in the arginine biosynthesis pathway partial mRNA

ATGTCGAATCAAGAAGAAGGACTTAAACTGTGGGGTGGCAGGTTTACT
GGGGCTACTGACCCCTTGATGGATTTGTATAACGCTTCCTTACCTTACG
ACAAGAAAATGTACAAGGTGGATTTAGAAGGAACAAAAGTTTACACTG
AGGGCCTGGAGAAAATTAATTTGCTAACTAAAGACGAACTAAGTGAGA
TTCATCGTGGTCTCAAATTGATTGAAGCAGAGTGGGCAGAAGGGAAGT
TTGTTGAGAAGCCAGGGGATGAGGATATTCACACTGCTAATGAACGTC
GCTTGGGTGAGTTGATTGGTCGTGGAATCTCTGGTAAGGTTCATACCGG
AAGGTCTAGAAATGATCAAGTTGCCACTGATATGCGGTTGTATGTCAGA
GACAATCTAACTCAGTTGGCTGACTATCTGAAGCAGTTCATTCAAGTAA
TCATCAAGAGAGCTGAACAGGAAATAGACGTCTTGATGCCCGGTTATA
CTCACTTGCAAAGAGCTCAACCAATCAGATGGTCTCACTGGTTGAGCAT
GTATGCTACCTATTTCACTGAAGATTATGAGAGACTGAATCAAATCGTT
AAAAGGTTGAACAAATCCCCATTGGGAGCTGGAGCTTTGGCTGGTCATC
CTTATGGAATTGATCGTGAATACATTGCTGAGAGATTAGGGTTTGATTC
TGTTATTGGTAATTCTTTGGCCGCTGTTTCAGACAGAGATTTTGTAGTCG

AAACCATGTTCTGGTCTTCGTTGTTTATGAATCATATTTCTCGATTCTCA
GAAGATTTGATCATTTACTCCACTGGAGAGTTTGGATTTATCAAGTTGG
CAGATGCTTATTCTACTGGATCTTCTCTGATGCCTACAAAAAAAAACCC
AGACTCTTTGGAGTTATTGAGGGGTAAATCTGGTAGATGTTTTGGGGCC
TTGGCTGGTTTCCTCATGTCTATTAAGTCCATTCCGTCAACCTATAACAA
AGATATGCAAGAGGATAAGGAGCCTTTATTTGATACTCTAATCACTGTA
GAGCACTCGATTTTGATAGCATCCGGTGTAGTTTCTACCTTGAACATTG
ATGCCGAACGAATGAAGAATGCTCTAACTATGGATATGCTGGCTACAG
ATCTTGCCGACTATTTAGTTAGAAGGGGAGTTCCATTCAGAGAAACTCA
CCACATTTCTGGTGAATGTGTCAGACAAGCCGAGGAGTTGAACCTTTCT
GGTATTGATCAGTTGTCCCTCGAACAATTGAAATCCATTGACTCCCGTT
TTGAGGCTGATGTGGCTTCAACGTTTGACTTTGAAGCCAGTGTTGAAAA
AAGAACTGCCACCGGAGGAACTTCTAAGACTGCTGTTTTAAAGCAATTG
GATGCACTGAATGAAAAGCTAGAGTCTTGA

PYC2

cbs7435_chr2 cbs7435_chr2:46620..50144 (+ strand) class=gene length=3525

ATGGCCGAAGAAGACTACTCCCCGCTGTACCAGCTGAGACGTGACTCG
TCTCTTCTGGGCACAATGAACAAGATTTTGGTTGCCAACAGAGGCGAAA
TTCCAATTCGTATCTTTAGAACAGCCCATGAGCTCTCAATGAACACGGT
GGCCATCTATTCTCACGAGGATCGTTTGTCGATGCACAGATTGAAGGCT
GACGAAGCTTATGTCATCGGTGAGCGTGGCCAATACTCTCCAGTCCAGG
CTTATCTTGCAATTGACGAGATCATCAAGATTGCCGTGAAGCACAACGT
TAACATGATTCACCCAGGTTACGGATTTTTGTCTGAAAATTCGGAGTTT
GCTAGAAAAGTCGAGGAAAATGGTATCTTGTGGGTTGGACCGTCCGAT
ACCGTCATCGACGCTGTTGGAGACAAAGTTTCGGCCAGAAACTTGGCCT
ACGCTGCCAATGTTCCAACCGTCCCTGGTACCCCGGGTCCAATTGAAGA
CGTTGCTCAGGCTACAGCTTTTGTTGAGGAATACGGTTACCCAGTTATC
ATCAAGGCCGCCTTTGGGGGTGGTGGAAGAGGTATGCGTGTGGTTCGT
GAAGGAGATGACATAGAAGATGCTTTCCTTAGAGCTTCTAGTGAGGCT
AAGACTGCGTTCGGTAACGGTACTGTTTTCATTGAACGTTTCTTGGACA
AACCGAAGCACATTGAGGTCCAATTGTTGGCTGATAACTACGGTAATGT
TATTCATTTGTTTGAGAGAGATTGTTCTGTTCAAAGAAGACACCAAAAG
GTTGTCGAAATTGCCCCTGCAAAGACACTCCCAGTTGAGGTCCGCAATG
CAATCTTGAATGACGCTGTCAAATTAGCCAAGACTGCCAATTATAGAAA
CGCTGGTACCGCTGAGTTTTTGGTGGACAGCCAGTACAGACACTACTTT
ATTGAGATTAACCCTCGTATTCAAGTTGAGCACACCATTACTGAGGAGA
TCACTGGTGTGGATATAGTAGCTGCCCAAATTCAAATTGCTGCTGGTGC
TTCGTTAGAACAACTAGGCTTATTGCAAGAAAAAATCACCACACGTGG
GTTTGCTATTCAATGTCGTATCACCACTGAAGATCCTACGAAGAATTTC
CAACCAGACACGGGAAAGATCGAAGTTTATAGATCTTCTGGTGGTAAC
GGTGTCAGACTTGACGGTGGTAACGGGTTTGCTGGTGCCGTCATCTCTC
CCCATTATGATTCGATGTTGGTCAAGTGTTCCACTTCTGGCTCTAACTAC

GAGATTGCTCGTCGTAAGATGATCAGAGCTTTAGTTGAATTCAGAATTA
GAGGTGTTAAGACCAACATTCCTTTCTTGTTGGCTTTACTGACTCACCCT
GTTTTCATGACCTCCGAATGTTGGACTACATTTATTGATGATACCCCAG
AATTGTTCAAAATGTTGACTTCTCAGAACAGAGCTCAGAAGCTGCTGGC
TTACTTAGGTGATCTGGCTGTTAATGGTTCCTCAATCAAGGGACAATTG
GGTCTGCCAAAACTACATAAGGAAGCTGATATCCCCGCAATCACTGAC
ATAAATGGAGATGTCATTGATGTTTCTATTCCTCCTCCTGATGGATGGA
GACAGTTTTTGTTGGAGAAAGGTCCTGAACAATTTGCTCAACAAGTACG
TGCCTTCCCTGGTTGTATGATTATGGACACTACCTGGAGAGATGCTCAC
CAGTCTTTGCTCGCAACTCGTGTTAGAACTATTGACTTACTAAATATTGC
TCCAGCTACTTCTTACGCCCTGCATCACGCTTTTGCACTGGAATGTTGGG
GAGGTGCCACTTTTGATGTTTCCATGCGTTTCTTGCACGAGGACCCATG
GCAAAGATTGCGCAAACTCCGTAAAGCTGTACCAAACATCCCATTCTCC
ATGTTATTGAGAGGTGCTAACGGTGTTGCCTACTCTTCTCTCCCAGATA
ACGCTATTGATCACTTTGTTAAGCAGGCCAAAGACACTGGTGTTGATGT
CTTCCGTGTATTTGATGCTCTTAACGATATCGAGCAATTGAAGGTCGGA
GTTGACGCTGTCAAGAAAGCTGGAGGTGTTGTTGAAGCCACCATGTGTT
ACTCCGGAGACATGTTGAAACCAGGTAAAAAGTATAACTTGGAATACT
ACATCAACCTGGCCACTGAGATTGTTGAGATGGGTACCCACATATTAGC
TGTAAAAGACATGGCTGGTACGTTGAAACCTACTGCTGCCAAACAGTTA
ATCTCTGCTTTGAGGAGAAAGTTCCCCTCGTTGCCAATCCACGTTCATA
CCCATGACTCTGCCGGTACTGGTGTTGCCTCCATGGTGGCCTGTGCTAG
AGCTGGTGCTGATGTTGTGGACTGTGCTGTTAACTCCATGTCTGGTATG
ACTTCTCAACCTTCAATGTCTGCTTTCATTGCGTCTTTAGATGGAGAAAT
TGAGACTGGTATCCCTGAAGCTAATGCTAGAGAGATTGATGCTTACTGG
GCCGAGATGAGACTTTTGTACTCATGTTTCGAGGCTGATCTTAAGGGTC
CTGACCCTGAAGTCTACCAGCATGAAATCCCAGGTGGTCAATTGACTAA
CTTACTCTTCCAAGCTCAGCAAGTTGGTCTGGGAGAAAAATGGGTTGAA
ACTAAGAAGGCATATGAGGCAGCTAACAGGTTACTCGGTGACATTGTA
AAGGTTACGCCAACCTCTAAGGTCGTTGGTGACTTGGCTCAATTCATGG
TATCTAATAAATTGTCATCCGAAGATGTCGAAAGATTGGCATCTGAGTT
AGATTTCCCAGATTCTGTTCTTGACTTTTTTGAAGGCTTAATGGGTACTC
CATATGGAGGTTTCCCTGAACCATTGAGAACAAATGTCATCTCAGGTAA
AAGAAGAAAGCTAACTAGCAGACCTGGTTTGACTTTGGAGCCATACAA
CATTCCTGCCATTAGAGAGGATCTTGAGGCTCGTTTCAGTAAGGTTACT
GAGAATGACGTTGCCTCTTACAACATGTACCCAAAGGTTTACGAAGCCT
ACAAGAAGCAACAAGAATTATATGGTGATCTGTCTGTGCTTCCAACCAG
ACATTTCTTATCTCCTCCTAAAATTGATGAAGAGATACATGTTACCATT
GAGCAAGGAAAAACTCTAATCATTAAGTGTATGGCTGTTGGTGAGCTAT
CTCAATCATCTGGTACCAGAGAAGTCTACTTTGAGTTGAACGGTGAGAT
GAGAAAGGTCACAGTTGAGGACAAGAATGCCGCTGTTGAAACTATCAC
TAGACCTAAGGCTGATGCTCACAACCCTAACGAGATTGGTGCTCCAATG
GCTGGTGTCGTTGTCGAGGTTAGAGTGCACGAAAATGGCGAAGTCAAG
AAGGGTGATCCAATTGCAGTTCTATCTGCTATGAAGATGGAGATGGTTA
TCTCCTCCCCTGTTGCTGGTCGTATTGGACAGATCGCAGTCAAGGAGAA

CGACTCTGTTGATGCAAGCGACCTGATCGCCAAAATCGTCAAGGCTGA
GTAA

PDA1

>cbs7435_chr2 cbs7435_chr2:555932..557122 (+ strand) class=gene length=1191

ATGCTCAAGTTTCTCAGCCCTCAATCTAGAATTGCTGCAAGAGTGGCCA
ATGCCAGATATATGGCTTCTGCAAGCCCATCGACTGTTTCCATTGACCT
CCCAGCCTCTTCTTTCGAGACTTATGAGTTAGAGCAAGGTCCAGAGCTT
CAGTTTGAAACGGAGAAGGAGACTCTGCTGCAAATGTATAAGCAAATG
GTGATTGTCAGACGTATGGAAATGGCTTCAGATGCCCTCTACAAGGCTA
AAAAGATTAGAGGTTTCTGCCACTTGTCTGTAGGCCAAGAAGCCGTGGC
TGTAGGAATTGAGAGTGCTATAACTAAGAAGGACACTGTTATTACATCC
TATAGATGTCACGGTTTCACATATTTAAGAGGTGCCTCTGTCAAGGAAG
TCCAGGGTGAATTGATGGGTAAGAGATGTGGTGTGTCTTACGGTAAAG
GTGGTTCGATGCATATGTTCACTACTGGATTCTACGGTGGAAACGGAAT
TGTTGGTGCTCAAGTTCCAGTTGGTGCTGGTCTGGCCTTTGCACATCAGT
ACAGGAATGAGAAGAACTGTACATTTGCCCTGTACGGGGATGGTGCTT
CTAACCAAGGTCAAGTATTCGAGTCCTTTAATATGGCCAAGTTGTGGGA
TTTGCCTGTGATTTTTGCCTGTGAAAACAACAAGTATGGTATGGGAACT
TCTGCTTCTAGATCATCTGCTATGACTGAGTACTACAAGAGAGGGCAAT
TTATCCCCGGTTTGAAAGTGAATGGTATGGATGTTCTTGCTTGTTACCA
AGCTTCTAAGTTTGCCAAGGACTGGACTGTCTCGGGAAATGGTCCTCTA
GTAATGGAGTACGAGACCTACAGGTATGGTGGACACTCTATGTCCGAC
CCAGGAACCACTTACAGAACACGTGAAGAAGTGCAGAACATGAGATCC
AGAAATGACCCAATTGCAGGCTTGAAGATGCACTTAATTGAGCTTGGA
ATATCTACAGAGGAGGAACTCAAGGCCTATGACAAGGAGGCTAGAAAG
TATGTTGACAAGCAAACCAAAGAAGCTGAACTCGCTCCTCCCCCTGAA
GCCAAAATGGATATTTTATTTGAAGATGTTTATGTCAAGGGTTCCGAAC
CCCCCGTCCTGAGAGGAAGAATAAATGAGGATTCTTGGTCCTTCGAGA
AGAACGGCTTCGCTAATCGTTAG

PGK1

>cbs7435_chr1 cbs7435_chr1:1942992..1944242 (+ strand) class=gene length=1251

ATGTCTCTTTCAAATAAACTTTCAGTCAAAGATCTCGATGTTGCCGGAA
AGCGTGTCTTCATCCGTGTCGACTTCAACGTTCCTCTGGATGGTGACAA
GATCACCAACAACCAGCGTATCGTTGCTGCTTTGCCAACTATCCAATAT
GTTTTGGATCACAAGCCAAAGGTCGTCGTTCTGGCTTCTCATTTAGGCC
GTCCAAACGGAGAGGTCAACCCAAAATTCTCTTTAAAACCAGTTGCTGC
TGAATTGTCCTCCCTACTAGGTAAGAAGGTGACTTTCTTGAACGATAGT
GTTGGACCAGAGGTTGAGAAGGCTGTCAACTCTGCCTCCAATGGAGAG

GTTATTCTTTTGGAGAACTTGCGTTTCCACATTGAAGAAGAAGGATCTC
AAAAGAAAGATGGTCAAAAGATCAAGGCCGACAAGGAGGCTGTTGCC
AGGTTCAGAAAGCAATTGACCGCATTGGCCGATGTCTACGTTAACGAC
GCCTTCGGTACCGCTCACAGAGCCCACTCCTCCATGGTTGGATTTGAAT
TGGAGCAAAGAGCTGCTGGTTTCTTGATGGCTAAGGAGTTGACATACTT
CGCTAAGGCCCTGGAAAACCCTGTCAGACCATTCTTGGCCATCCTTGGT
GGTGCTAAGGTTTCTGACAAGATTCAATTGATTGACAATTTGCTGGACA
AGGTCGATTCCATCATCATTGGTGGAGGAATGGCTTTCACTTTTATCAA
GGTTTTGGATAACGTTGCCATTGGTAACTCTTTGTTCGACGAGGCTGGT
GCCAAGTTAGTTCCCGGCTTAGTTGAGAAAGCCAAGAAGAACAATGTC
AAACTGGTTCTTCCAGTCGACTTCGTCACTGCCGACGCCTTCTCCAAGG
ATGCAAAGGTCGGTGAAGCCACGGTTGAGTCTGGTATTCCAGACGGAT
TGCAAGGATTGGACGCTGGTCCAAAATCCAGAGAATTGTTCGCAGCTA
CCATCGCTGAGGCTAAGACAATCGTCTGGAACGGTCCTCCAGGTGTTTT
CGAGTTTGACAAGTTTGCTGAAGGTACCAAGTCTATGTTGGCAGCTGCC
ATCAAGAACGCTCAGAACGGTGGAACTGTCATCGTTGGTGGTGGTGAC
ACGGCTACCGTTGCTAAGAAGTTCGGTGGTGCTGACAAGCTATCCCACG
TTTCCACTGGAGGAGGAGCTTCTTTGGAACTGTTGGAGGGAAAGGAGC
TTCCAGGTGTAGTTTACTTGTCCAACAAGGCTTAA

ALD4

>cbs7435_chr2 cbs7435_chr2:1464707..1466284 (- strand) class=gene length=1578

ATGCTTAGAACTTCTCCAGCTACTAAGAAAGCTCTCAAGTCGCAGATTA
ACGCCTTCAACGTTGCTGCCTTGAGATTCTACTCCTCATTGCCTTTGCAG
GTTCCAATTACCTTGCCAAACGGTAAGACCTACAATCAGCCAACAGGTT
TGTTTATCAACAATGAGTTCGTTCCTTCTAAGCAAGGTAAGACCTTTGC
TGTTTTAAACCCTTCCACTGAGGAGGAGATTACTCACGTCTACGAGTCC
AGAGAGGACGACGTTGAGTTAGCCGTTGCAGCCGCTCAAAAGGCTTTC
GACTCAACCTGGTCCACCCAGGACCCTGCTGAGAGAGGTAAGGTCTTG
AACAAGTTGGCTGACCTGATCGAGGAGCACTCTGAGACCCTTGCCGCC
ATCGAGTCCTTGGACAACGGTAAGGCCATTTCCTCCGCTAGAGGTGATG
TTGGTCTGGTTGTCGCCTACTTGAAGTCCTGTGCCGGTTGGGCCGACAA
GGTTTTCGGTAGAGTTGTTGAAACCGGAAGCTCCCACTTCAACTACGTT
AGAAGAGAGCCATTGGGTGTTTGTGGTCAGATTATCCCATGGAACTTTC
CTCTTCTGATGTGGTCCTGGAAAGTTGGTCCAGCTTTGGCCACTGGTAA
CACTGTTGTCCTGAAGACAGCCGAGTCTACTCCTCTGTCCGCCCTGTAC
GTTTCCCAATTGGTCAAGGAGGCCGGTATCCCAGCTGGTGTCCACAACA
TTGTGTCCGGTTTCGGTAAGATTACTGGTGAAGCTATTGCTACTCATCCT
AAGATCAAGAAGGTTGCCTTCACTGGTTCTACCGCCACTGGTCGTCACA
TCATGAAGGCTGCTGCCGAATCCAACTTGAAGAAGGTTACTTTGGAGTT
GGGTGGTAAATCTCCTAACATCGTGTTCAACGATGCTAACATTAAGCAA
GCTGTCGCCAACATCATCCTCGGTATTTACTACAACTCTGGAGAAGTTT
GTTGTGCTGGTTCCAGAGTTTATGTTCAATCCGGTATTTACGACGAGCTT

TTGGCCGAATTCAAGACTGCTGCTGAGAATGTCAAGGTTGGTAACCCAT
TCGACGAGGACACCTTCCAAGGTGCTCAAACCTCTCAGCAACAATTGG
AGAAGATTTTGGGTTTCGTTGAGCGTGGTAAGAAGGACGGTGCTACTTT
GATTACTGGTGGTGGCAGATTAGGTGACAAGGGTTACTTCGTCCAGCCA
ACTATCTTCGGTGATGTTACACCAGAGATGGAGATTGTCAAGGAAGAG
ATCTTTGGTCCTGTTGTCACTATCAGCAAGTTTGACACCATTGATGAGG
TTGTCGACCTTGCTAACGACTCTCAATACGGTCTTGCTGCTGGTATCCAC
TCTGACGATATCAACAAGGTCATTGACGTTGCTGCTAGAATCAAGTCCG
GTACCGTGTGGGTCAACACCTACAACGATTTCCACCAAATGGTTCCATT
CGGTGGATTTGGCCAATCCGGTATTGGTCGTGAGATGGGTGTTGAAGCT
TTGGAAAACTACACCCAATACAAGGCTATCCGTGTCAAGATCAACCAC
AAGAACGAGTAA

FBP1

>cbs7435_chr3 cbs7435_chr3:551377..552438 (+ strand) class=gene length=1062

ATGTCCAATAACACCACCCAAAACTTGGCTGAACAGAAGGGTATTCAG
ACCGACTTGGTCACTCTGACTCGTTTCATCTTGGACGAGCAGAAGAAAT
CTGCCCCAAACGCTACCGGTGAGCTGACTTTGCTTCTGAACTCCCTGCA
ATTTGCCTTCAAATTCATTGCTCACACCATCAGAAGATCCGAATTGGTC
AACTTAATTGGATTGGCCGGTGTGACCAACGCCACTGGTGACGATCAG
AAAAAGCTGGATGTTATTGGAGATGAGATCTTCATCAATGCCATGAAG
GGATCCGGTAACGTTAAGTTACTTGTTTCCGAAGAGCAAGAAGACCTTA
TTGTGTTTGAGTCCTCCAAGGGAAACTATGCTGTTGTCTGCGACCCAAT
TGACGGTTCATCCAACTTGGACGCTGGTGTCTCAGTTGGTACCATTTTTG
GTGTATACAAACTGTTGCCAGGCTCAGCTGGTTCTATTAAGGATGTTTT
GAGATCAGGAACAGAGATGGTGGCTGCCGGTTACACCATGTACGGTGC
ATCTTCTCACTTGATGCTTACTACCGGAAATGGAGTCAACGGTTTCACC
TTGGACACCGACTTGGGAGAGTTCATTCTCACGTATCCATCCTTAAAAA
TTCCGCACACCAGGGCTATCTACTCTATCAATGAAGGTAACTCGCACTA
CTGGACCGATGGAGTCAATGAGTACATTGCTTCTTTGAAGAAACCCCAA
GCAAACGGAAAACCATACAGTGCCCGTTATATCGGATCAATGGTTGCT
GATGTCCACAGAACTCTGTTATATGGAGGTATCTTTGGCTATCCAGCAG
ACTCCAAGTCTAAGTCCGGTAAACTGAGAGTCTTGTACGAGTGTTTCCC
AATGGCCCTCCTGTTGGAGCAGGCTGGTGGTGAGGCAGTCAACGATAA
GGGTGAGAGAATTCTAAACTTGGAGCCAAAACAAGTGCATGAGAGATC
TGGTATCTGGTTGGGTTCCAAGGGAGAAGTAGAGAGATTACTACCCTAT
CTAACAAAAAAAATTAAGATCCAATCTGTGAATCTCTAG

MLS1

>cbs7435_chr4 cbs7435_chr4:1395371..1397053 (+ strand) class=gene length=1683

ATGGCTGCCCCAAGAACTTACGGCTCCCTTGAAGGAGTGCAATTGTATG
GAAAAGTTGACTCAACCCCACTGTATGATAGTCCAATTACCCCTGCTGA
TATCCTGACCAAAGATGCTTTGAAATTTGTCGTTCTACTTCACAGAACTT
TCAACAGTACCAGAAAGGAGCTTCTTGAGAATAGACAGAAGGTCCAGG
CCAAATTGGACAAAGGTGAGAAGTTGACTTTCCTCCCCGAGACTAAGT
ACATCAGGGATGATCCTAACTGGAAGTGTAAGGCTCCAGCACCAGGAC
TGACAGATCGTCGTGCCGAGATCACTGGTCCTCCAGAGAGAAAGATGG
TCGTCAATGCCTTGAACACAGATGTTTATACCTACATGACTGATTTTGA
GGACTCCTGCTCTCCAACATGGTCTAATATTATTTACGGACAAGTTAAT
TTGTATGATGCTCTCAGAGATCAAATTGACTTCACTTTGCCTGCCACTG
GCAAATCATACAAGGTCAAGAAGGAAGGTCGCAACCCACCAACCATCA
TTGTCAGACCTCGTGGATGGCATATGGTTGAGAGCCATGTCACCGTTGA
CGGTGAGCCAATTAGTGCCTCTCTATTCGACTTTGGTCTTTACTTCTTCC
ACAATGCCCGTCAATTGCTGGACAATGGTGTTGGTCCCTACTTTTATTTG
CCAAAGATGGAGCACTGGTTAGAGGCAAAGTTGTGGAATGACGTTTTC
AACGTCGCTCAAGACGCTTTGGCCATTCCTCGTGGTACCATCAGTGCTA
CTGTCTTGATCGAGACTTTGCCAATTTCCTACCAATTGAACGAAGTTTTG
TATGCCCTAAGAGATCATTCCGCTGGTTTGAACTGTGGACGTTGGGATT
ACATGTTCTCCACTATCAAACGTCTACGTAATGACCCCAAGCACATTTT
GCCAGACCGTGGTCTTGTGACCATGAAGGTTCCTTTCATGACCAACTAT
GTTAAGCAGTTAATTAAGGTCTGTCATCAGCGTGGAGTACACGCTATGG
GAGGAATGGCTGCCACAATTCCTATTAAGAACGATGCTGAACGTAATG
CCAAGGCCATGCAAGCTGTTCATGATGACAAACTACGAGAAGTTCTTGC
TGGCCACGACGGTACCTGGATTGCCCATCCAGCTTTGGCACCAATCGCT
CTTGAAGTCTTCAACAAGCACATGCCAACTCCTAACCAATTGCACAAGA
TTCCCGTCTACGACAGGGAAGTTACTGAGGAAGATCTGGTTGATACCAA
CATCGATGGTTTCAAGATCACTAAAGACGGTATTCTGATTAACATTTAT
ATCGGTCTGAACTATATGGAGGCTTGGTTGCAAGGTTCTGGATGTGTGC
CTATCAATCACTTGATGGAAGATGCTGCCACTGCTGAAGTTTCTCGACT
GCAATTATTCTCCTGGGTCAAGCATGGTGTAAAACTTACTGACACTGGC
GAGAAGATCACAAAGGACCTGGTTGTCAAGTTGATCGACGACGAAGTT
GCAAAGCTGAGCCCCAGCAGACCCAACAACAAGTTTGACATTGCTGCC
GACTGTTTGAAGAAGGAGATTAGTGGAGAAGTCGAGGTTGCTGAGTTT
TTGACCGACTTACTGTACCCAGACGTTGTTACTCTTGAGTCTTCTCCTGT
TGATTTGGACAGTCTTAAATAG

DAS1

>cbs7435_chr3 cbs7435_chr3:634689..636812 (+ strand) class=gene length=2124

ATGGCTAGAATTCCCAAAGCAGTTTCTTACAATGATGACATCCATGACT
TGGTCATCAAAACCTTCCGTTGTTACGTTCTCGACTTAGTCGAACAGTA
TGGTGGTGGTCACCCTGGTTCTGCCATGGGTATGGTCGCCATTGGTATC
GCTCTGTGGAAGTACCAGATGAAGTACGCTCCAAATGATCCAGACTACT
TCAACAGAGATCGTTTTGTCTTGTCAAACGGTCACGTCTGTTTGTTCCAA
TACTTGTTCCAGCACTTAACTGGTTTGAAGGAGATGACTGTCAAGCAAC
TTCAATCTTACCACTCTTCCGATTATCACTCATTGACTCCTGGACACCCT
GAAATTGAGAACCCTGCTGTTGAGGTTACCACTGGTCCCCTGGGACAAG
GTATCTCTAACGCTGTCGGTATGGCCATTGGTTCAAAGAACCTGGCCGC
TACTTACAACAGACCTGGCTTCCCTGTCGTTGACAACACTATCTATGCT
ATTGTTGGTGATGCTTGTTTGCAAGAGGGACCTGCTTTGGAATCGATTT
CCTTAGCCGGTCACTTGGCCTTGGACAACCTTATTGTGATCTACGACAA
CAACCAGGTTTGTTGTGATGGTTCCGTCGATGTTAACAACACCGAAGAC
ATCTCCGCAAAGTTCAGAGCTCAGAACTGGAATGTTATCGACATTGTAG
ACGGTTCTAGAGATGTCGCTACCATTGTCAAGGCTATCGATTGGGCCAA
GGCTGAGACTGAGAGACCAACTCTGATCAACGTTAGAACTGAAATTGG
ACAGGATTCTGCTTTCGGTAACCACCACGCTGCTCACGGTTCTGCTCTA
GGTGAGGAAGGTATCCGGGAGTTGAAGACTAAGTACGGTTTTAACCCT
GCCCAAAAGTTCTGGTTCCCTAAAGAAGTATACGACTTCTTTGCTGAGA
AACCAGCTAAAGGTGACGAGTTAGTAAAGAACTGGAAAAAGTTAGTTG
ATAGCTATGTCAAAGAGTACCCTCGTGAGGGACAAGAGTTCCTTTCTCG
TGTTAGAGGTGAGCTTCCAAAGAACTGGAGAACTTACATTCCTCAAGAC
AAGCCTACCGAACCAACCGCCACCAGAACCTCTGCTAGAGAAATTGTT
AGGGCCCTTGGAAAGAACCTTCCTCAAGTTATTGCCGGTTCCGGTGACT
TATCTGTCTCAATTCTTTTGAACTGGGACGGAGTGAAGTACTTCTTCAA
CCCTAAGTTACAGACTTTCTGTGGATTAGGTGGTGACTACTCTGGTAGA
TATATTGAGTTTGGTATCAGAGAACACTCTATGTGTGCTATTGCCAACG
GTTTGGCTGCATACAACAAGGGTACTTTCTTGCCTATTACCTCTACCTTC
TACATGTTCTACCTGTATGCAGCACCTGCCTTGCGTATGGCTGCTCTTCA
AGAGTTGAAAGCGATTCACATTGCTACACACGACTCTATTGGAGCTGGT
GAAGATGGTCCAACCCACCAGCCTATTGCTTTGTCTTCATTATTCAGAG
CTATGCCCAACTTCTACTACATGAGACCAGCCGATGCTACCGAAGTTGC
AGCTCTGTTTGAAGTGGCTGTTGAGCTTGAACACTCCACATTGCTTTCTC
TGTCCAGACACGAGGTTGACCAATACCCAGGTAAGACTTCTGCCCAAG
GAGCCAAAAGAGGTGGTTACGTTGTTGAAGACTGCGAAGGAAAGCCAG
ATGTGCAACTGATCGGAACTGGTTCCGAGTTGGAATTCGCTATTAAGAC
TGCTCGTTTGCTAAGACAACAGAAGGGATGGAAGGTCAGAGTTCTGTC
ATTCCCATGTCAGAGATTGTTTGACGAGCAGTCTATTACTTACAGACGT
TCCGTCCTTAGAAGAGGAGAAGTTCCAACTGTCGTTGTTGAGGCCTATG
TCGCATACGGATGGGAGAGATACGCCACTGCTGGTTACACCATGAACA
CCTTCGGTAAGTCTCTTCCTGTTGAGGATGTCTACAAATACTTCGGATA
CACTCCTGAGAAGATTGGTGAGAGAGTGGTTCAATATGTCAACTCTATC
AAGGCTAGTCCTCAAATCCTTTACGAATTCCACGACTTGAAGGGAAAAC
CAAAGCATGACAAGTTGTAA

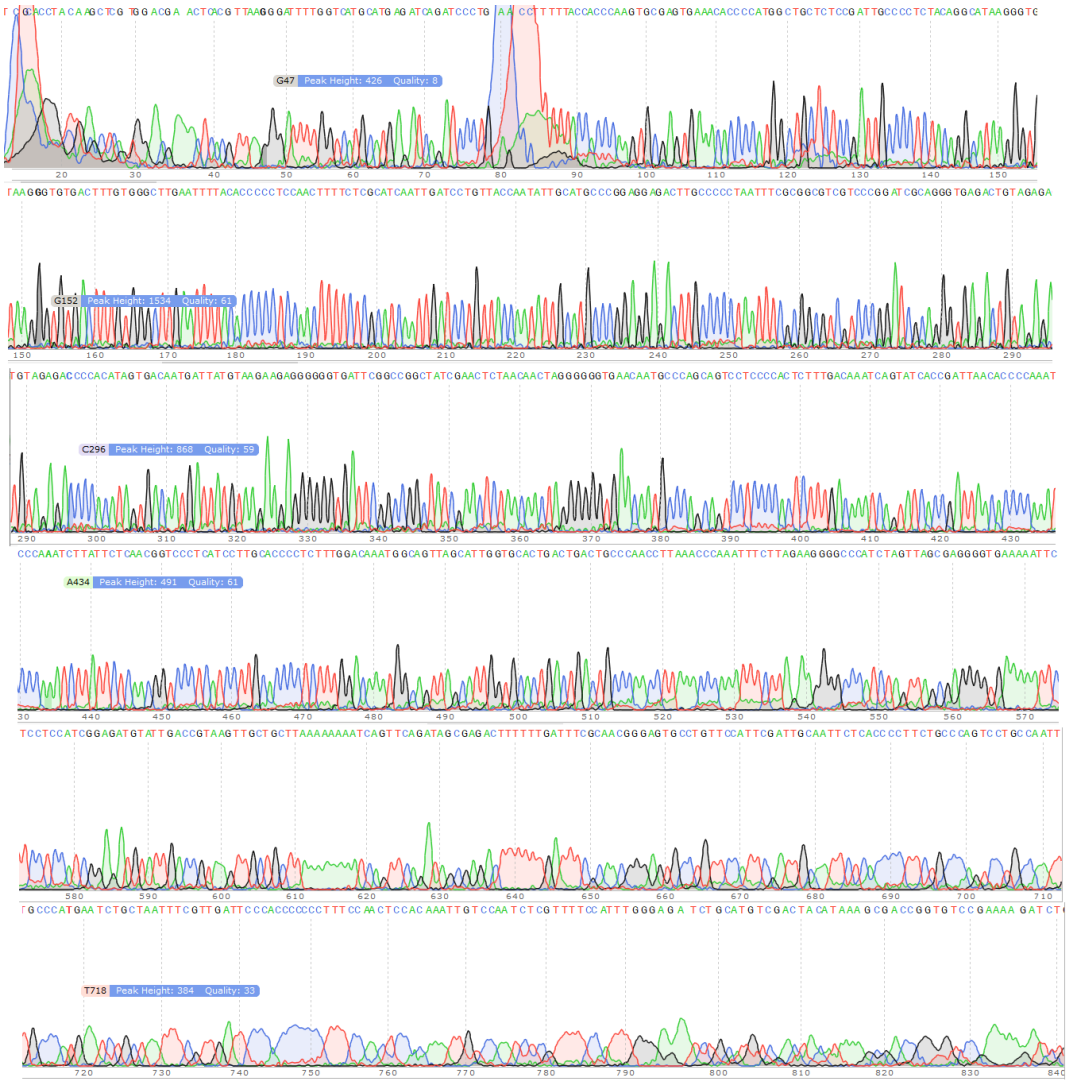
ACT1

>cbs7435_chr3 cbs7435_chr3:1776679..1778406 (- strand) class=gene length=1728

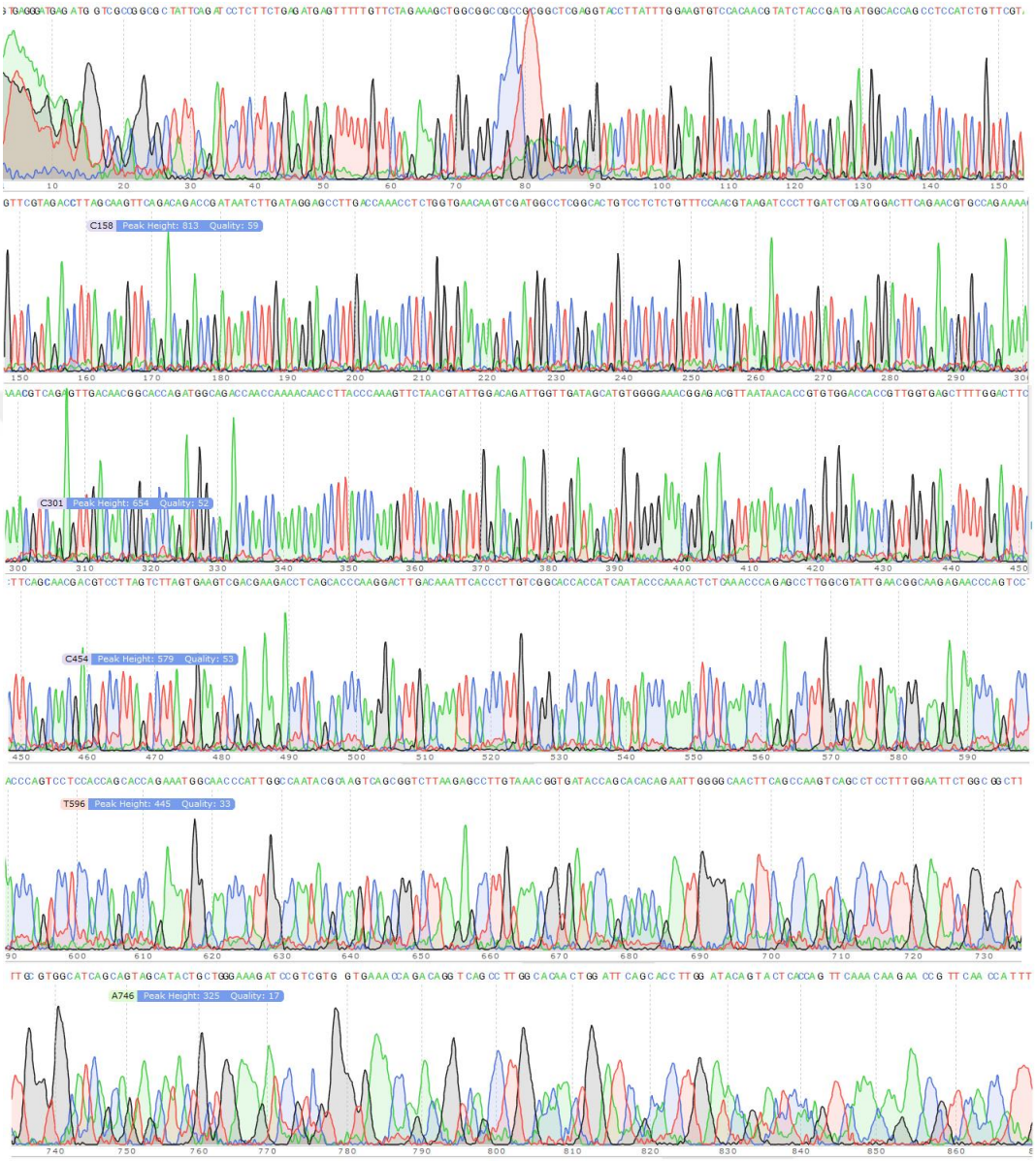
ATGGACGGAGGTAAGTCAGCTCTTTCAGCAAGGGAGCATGGAATCAGA
TTGGCAACGACTTCTGCGATTCTTATTCTTCCGATTGATTCAACAGTAGA
TTTGCGAGTGGTTGGCCAAACACACAGTTTCGATGCGATCTCCGTTTGA
TTTGTATGAATACGCAGTCATCTCATTGGTGGAGCATTTTATACGGATT
GATTGAGTGGAAGCATGGTACTGAGAGTATCTACATGTGCATTTCTGTT
GGTTGTCCTCGGTTGAGTTTTCGTATCTGTTTGTCCCCCCAATTTTGTCC
ACTCCTTTTGGCAATCTGCCGAATCCAAGCTCATAATACTAACTCCCAT
TTTAGAAGACGTTGCAGCAGTAAGTACCCTTACATGTAATCCATGAAAG
CAAATGTGGCGCAGTCAATCGAAGTCACTGTCTGATTGATTCTTCCCTC
AGTGGCGGTTAGGATTTTTCAATGGTAGCGCGCCCATTTCTTACCGTTG
GGAAGGGATTAGAATTTGATTTCGTAGCAGGATTCTGTGTTGACTGTGC
AGCCTCGCGCTTGTTTTTTCATCTTTCATTTCCTAATGAGGCAATATTGT
ATCATGATAGTATACTAACACAGACTTTTAGTTAGTTATCGACAATGGT
TCCGGTATGTGTAAGGCCGGATACGCCGGAGACGACGCCCCACACACA
GTGTTCCCATCGGTCGTAGGTAGACCAAGACACCAAGGTGTCATGGTCG
GTATGGGTCAAAAGGACTCCTTCGTCGGTGACGAGGCTCAATCCAAGA
GAGGTATCTTGACCTTGAGATACCCAATCGAGCACGGTATCGTCACTAA
CTGGGACGATATGGAAAAGATCTGGCACCACACCTTCTACAACGAGTT
GCGTCTGGCCCCAGAAGAGCACCCAGTTCTTTTGACTGAGGCTCCAATG
AACCCAAAGTCCAACAGAGAGAAGATGACCCAAATCATGTTCGAGACT
TTCAACGTTCCAGCCTTCTACGTTTCTATTCAGGCCGTTTTGTCCCTGTA
CGCTTCCGGTAGAACCACTGGTATCGTTTTGGACTCTGGTGACGGTGTT
ACCCACGTTGTCCCAATTTATGCCGGTTTCTCCTTACCACACGCTATTTT
GCGTATCGACTTGGCCGGTAGAGATTTGACCGACTACTTGATGAAGATC
TTGTCTGAGCGTGGTTACACTTTTTCTACCTCTGCTGAGAGAGAAATCG
TCCGTGACATCAAGGAGAAGCTTTGTTACGTTGCTCTTGACTTTGACCA
GGAATTGCAAACTTCTTCTCAATCTTCATCCATTGAGAAGTCTTACGAG
TTGCCAGATGGCCAAGTTATCACTATCGGTAACGAGAGATTCAGAGCTC
CTGAGGCTTTGTTCCACCCATCTGTACTTGGCCTTGAGGCTTCTGGTATC
GACCAAACCACTTACAACTCCATCATGAAGTGTGATGTTGATGTTCGTA
AGGAACTCTACAGTAACATCGTTATGTCCGGTGGTACTACTATGTTCCC
AGGTATTGCTGAGCGTATGCAAAAGGAGCTTACTGCCTTGGCTCCATCT
TCGATGAAGGTCAAGATTTCTGCTCCACCAGAAAGAAAGTACTCCGTAT
GGATCGGTGGTTCTATCCTCGCTTCTTTGGGTACTTTCCAACAAATGTGG
ATCTCAAAGCAAGAGTACGACGAATCTGGACCATCCAT

B. Nucleotide sequences of constructed plasmids

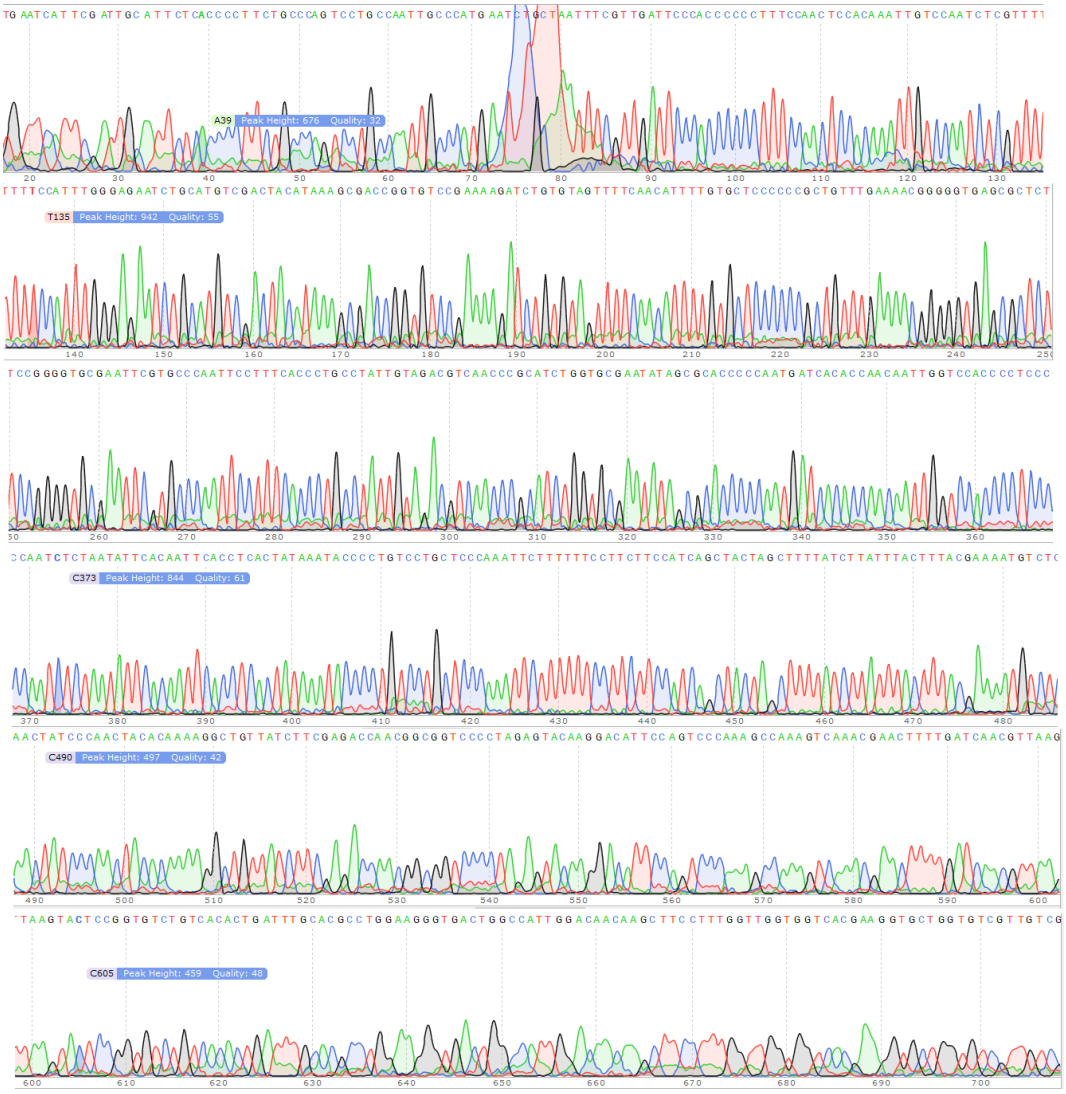
$P_{ADH2\text{-}wt}$- Forward



$P_{ADH2\text{-}wt}$- Middle

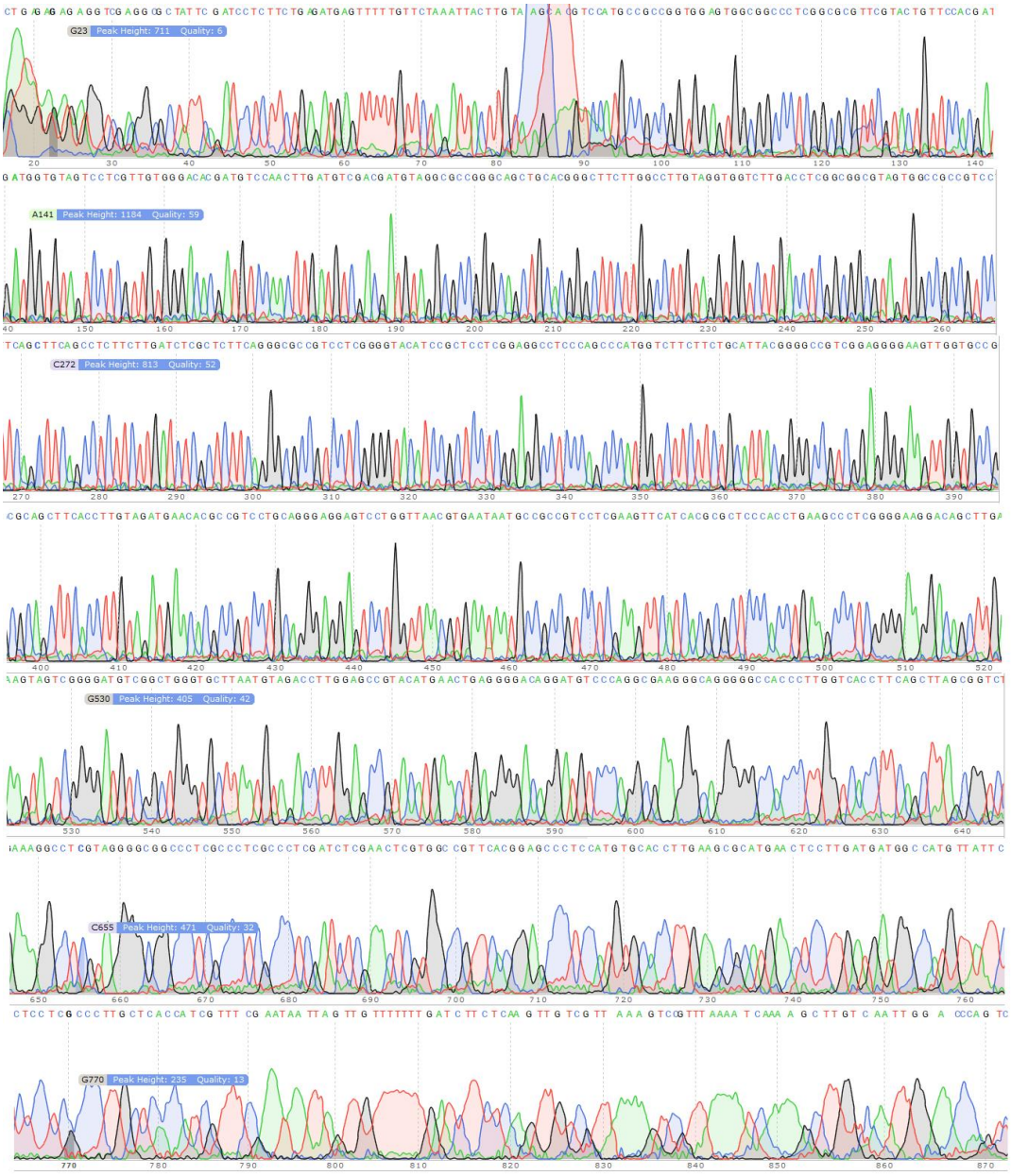


ADH2- Reverse



P_{mAOX}-Forward

mApple-Reverse



C. Buffers and Stock Solutions

1X TAE Agarose Gel Electrophoresis Buffer	40 mM Tris, 20 mM acetic acid, 1 mM EDTA.
1 M Tris-Cl (pH 8.0)	Dissolve 12.1 g Tris base 80 mL dH2O and adjust the pH to 8.0 by adding concentrated HCl. Make up the volume up to 100 mL. Autoclave and store at room temperature.
0.5 M Ethylendiaminetetra acetic acit EDTA (pH:8.0)	Dissolve 18.61 g EDTA in 80 ml distilled water. Adjust pH to 8.0 and bring final volume up to 100 mL. Autoclave and store at room temperature.
TE Buffer	Mix 1 mL 1M Tris-HCl (pH 8.0) and 0.5 mL of 0.5M EDTA, and complete to 100mL with dH_2O.

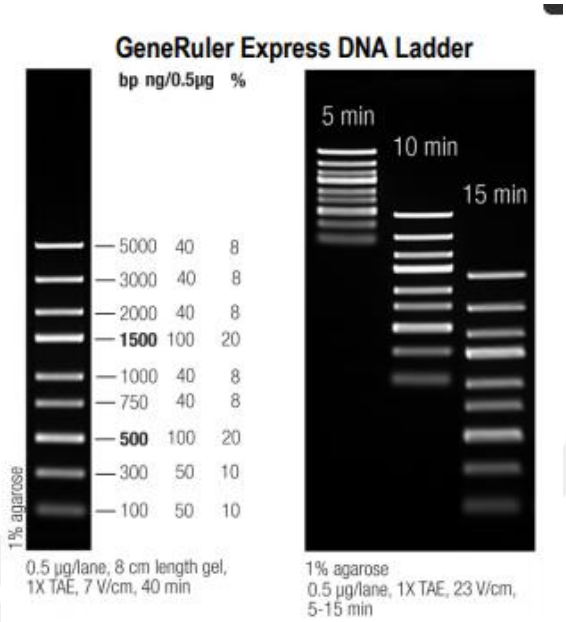
Calcium chloride Transformation Solutions

$MgCl_2$-$CaCl_2$ Solution	Dissolve 8.13 g $MgCl_2.6H_2O$ with 1.1099 g $CaCl_2$ in 500 mL dH_2O. Filter-sterilize and store at +4°C.
0.1 M $CaCl_2$ Solution	Dissolve 11.1 g $CaCl_2$ in 1 L dH_2O. Filter-sterilize and store at +4°C.

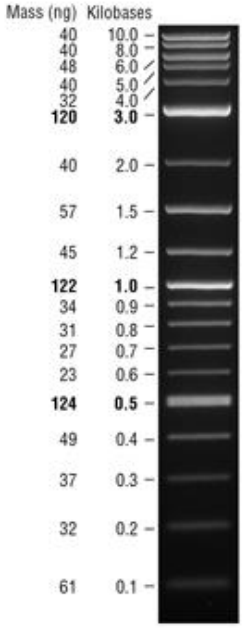
Lithium Chloride Transfection Solutions

1M LiCl	Dissolve 4.24 g of LiCl in distilled water and filter sterilize. Dilute with sterile water when needed.
PEG	Dissolve 50 % polyethylene glycol (PEG-3350) in distilled water and filter sterilize. Store in tightly capped bottle.
Single-stranded DNA	2 mg/mL denaturated, fragmented salmon sperm DNA in TE (pH 8.0) buffer, store at -20°C.

D. DNA Markers



Quick-Load® Purple 1 kb Plus DNA Ladder



E. Standard Curves for Organic Acids for HPLC Analysis

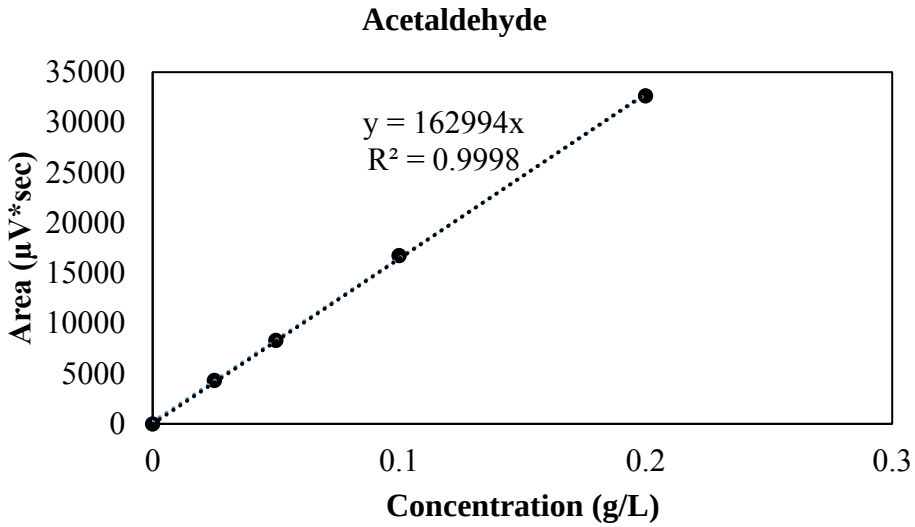


Figure E1 Standard curve for acetaldehyde

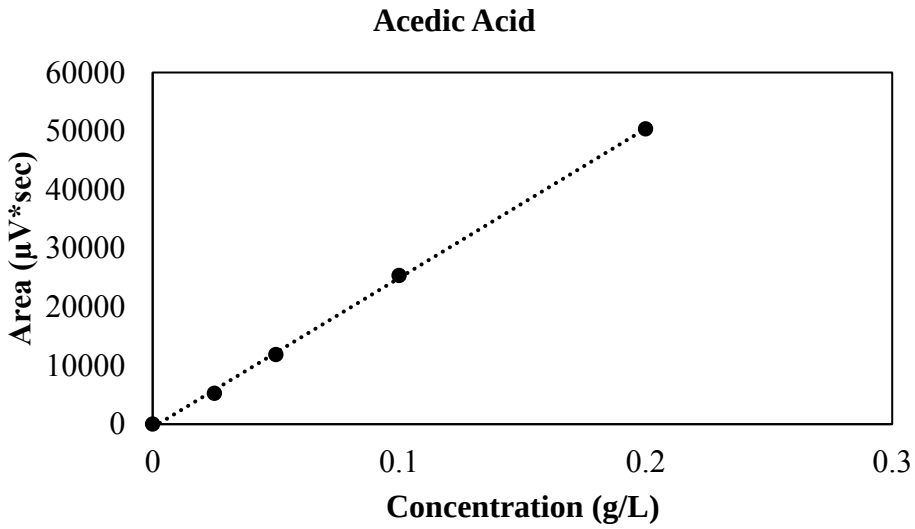


Figure E2 Standard curve for acetic acid

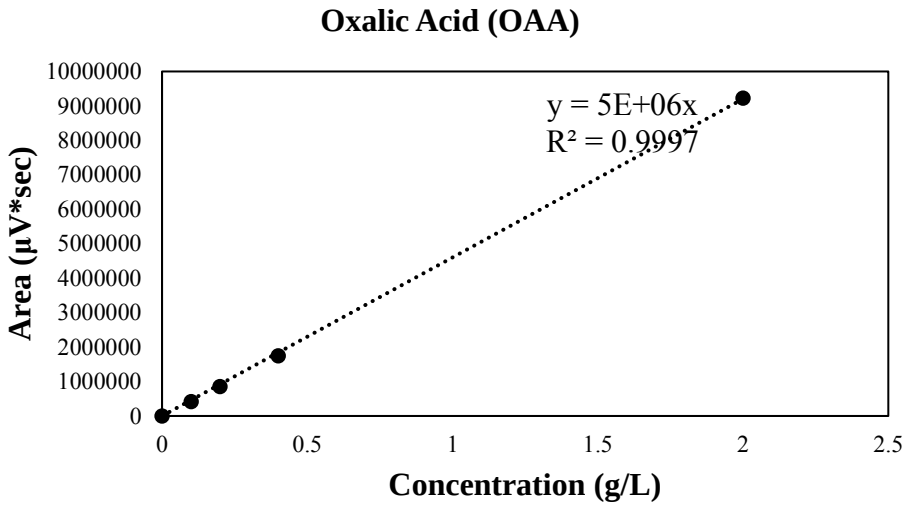


Figure E3 Standard curve for oxalic acid

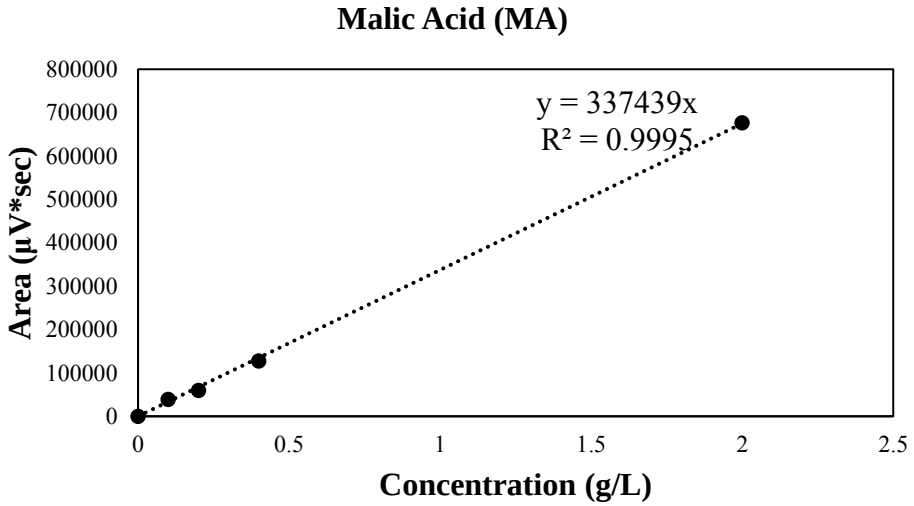


Figure E4 Standard curve for malic acid

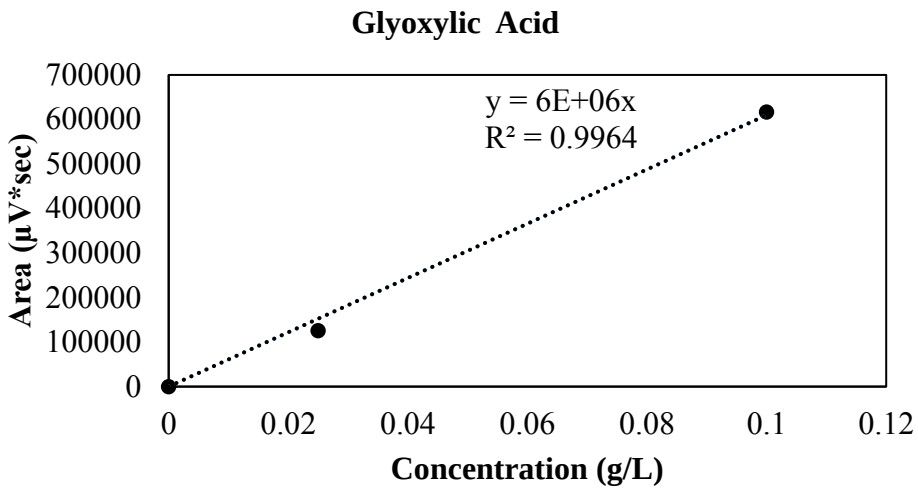


Figure E5 Standard curve for glyoxylic acid

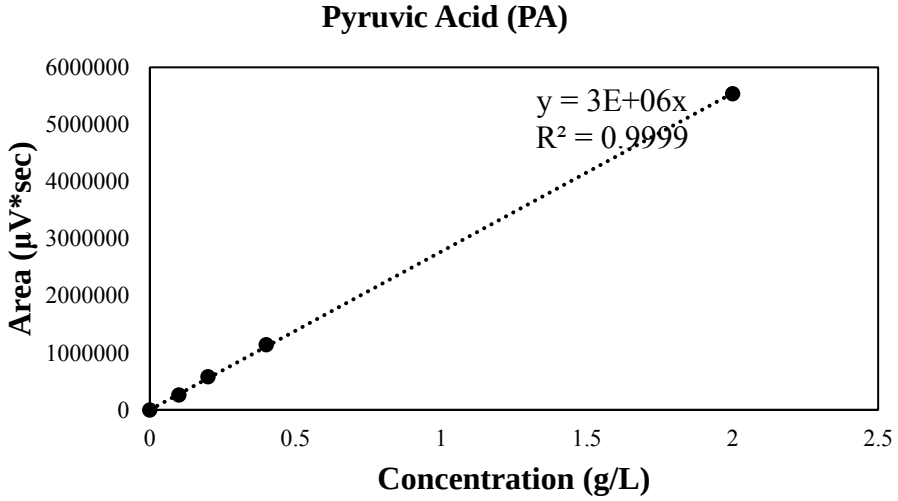


Figure E6 Standard curve for pyruvic aci