

PROTEIN ENGINEERING AND COVALENT MODIFICATION OF *Trichoderma reesei* CELLULASES in *Pichia pastoris* for TEXTILE BIOFINISHING

by  
GÜNSELİ BAYRAM AKÇAPINAR

Submitted to the Graduate School of Engineering and Natural Sciences  
in partial fulfillment of  
the requirements for the degree of  
Philosophy of Doctorate

Sabancı University  
March 2011

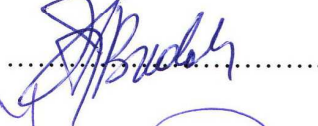
PROTEIN ENGINEERING AND COVALENT MODIFICATION OF *Trichoderma reesei* CELLULASES in *Pichia pastoris* for TEXTILE BIOFINISHING

APPROVED BY:

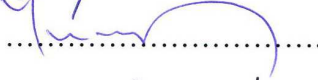
Assoc.Prof. Osman Uğur Sezerman  
(Dissertation Supervisor)



Assoc. Prof. Hikmet Budak



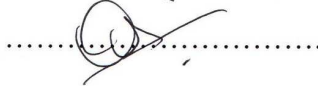
Prof. Yusuf Menceloğlu



Prof. Selim Çetiner



Prof. Pervin Anış



DATE OF APPROVAL: ....11.07.2011.....

© Günseli Bayram Akçapınar 2011

All Rights Reserved

*to my family and Tekin,*



## ABSTRACT

Cellulase enzymes have been extensively used for the biopolishing of cellulosic fabrics but they are inefficient to prevent pilling in viscose fabrics. Moreover, their application causes a loss in the fabric strength due to the aggressive action of the enzymes. One solution to this problem is the design and production of enzymes with increased molecular weights so that aggressive action of the cellulases would be limited to the fabric surface. In the framework of this study, cellulases and cellulase formulations that can ameliorate the problem of pilling and prevent loss of tensile strength in viscose fabrics were designed and produced. For this purpose, both protein engineering and chemical modification methods were used separately and in combination to obtain cellulases with desired properties. *Trichoderma reesei* Endoglucanase I (EGI), Endoglucanase III (EGIII), Cellobiohydrolase I (CBHI) enzymes were successfully cloned and expressed in *Pichia pastoris* under the control of AOX1 promoter to mg/L quantities. A loop mutant of EGI, (EGI\_L5) was prepared by introduction of a ten aminoacid long loop by molecular modelling and site directed mutagenesis for the creation of hotspots for directed crosslinking of the enzyme. The mutant enzyme was crosslinked using crosslinked enzyme aggregate (CLEA) technology. The effect of codon optimization on EGI production was analyzed. A mutant of EGI was prepared by inserting a second catalytic domain to EGI and thereby forming a bicatalytic mutant of EGI (EGI\_BC) with increased molecular weight. All of the recombinant enzymes were produced in a laboratory scale fermenter and characterized. A commercial cellulase preparation was crosslinked using CLEA technology and fractionated according to the particle size. The effects of native, engineered and chemically modified cellulases on viscose fabrics were evaluated. It was found that commercial cellulase preparation crosslinked using CLEA technology, recombinant EGI and EGI\_L5 produced in *P. pastoris* improved the pilling values of viscose fabrics by 20 % without much loss in the strength of the fabrics.

## ÖZET

Selülaz enzimleri selülozik kumaşların biyoparlatmasında yaygın olarak kullanılmaktadır ancak viskon kumaşlarda tüylülüğün önlenmesinde yetersiz kalmaktadırlar. Bu enzimlerin agresif hareketleri kumaş mukavemetinde kayıplara yol açmaktadır. Selülaz enzimlerinin aktivitesini kumaş yüzeyi ile sınırlayacak moleküler ağırlığı arttırılmış enzimler tasarlanarak ve üretilerek bu problem çözülebilir. Bu çalışmada, viskon kumaşlardaki tüylenme sorununu azaltabilecek ve kumaşların mukavemet kayıplarını engelleyebilecek selülaz ya da selülaz formülasyonları tasarlanmış ve üretilmiştir. Bu amaçla, hem protein mühendisliği yöntemleri hem de kimyasal modifikasyon yöntemleri istenilen özellikte selülazların elde edilebilmesi için hem ayrı ayrı hem de birarada kullanılmıştır. *Trichoderma reesei* Endoglukanaz I (EGI), Endoglukanaz III (EGIII), Sellobiyohidrolaz I (CBHI) enzimleri *Pichia pastoris*'te AOX1 promoter bölgesinin kontrolünde başarılı bir şekilde klonlanmış ve mg/L miktarlarında ifade edilmiştir. EGI'nin bir döngü mutanı olan EGI\_L5, 10 aminoasitlik bir döngünün moleküler modelleme ve yönlendirilmiş mutagenz yöntemleri kullanılarak enzimin yönlendirilmiş olarak çapraz bağlanması için sıcak noktalar oluşturmak üzere EGI'e eklenmesi ile oluşturulmuştur. Mutant enzim çapraz bağlı enzim agregatları (CLEA) teknolojisi kullanılarak çapraz bağlanmıştır. Kodon optimizasyonunun EGI üretimi üzerine etkisi araştırılmış ve kodon optimizasyonunun *P. pastoris*'te EGI üretimini % 24 arttırdığı saptanmıştır. EGI'e ikinci bir katalitik modül eklenerek böylece moleküler ağırlığı büyütülmüş EGI'in bikatalitik bir mutanı (EGI\_BC) elde edilmiştir. Tüm rekombinant enzimler laboratuvar ölçeğindeki bir fermentörde üretilmiş ve karakterize edilmiştir. Ticari bir selülaz formülasyonu CLEA teknolojisi kullanılarak çapraz bağlanmış ve parçacık büyüklüğüne göre parçalara ayrılmıştır. Ham, protein mühendisliği yoluyla değiştirilmiş ve kimyasal olarak değiştirilmiş selülazların viskon kumaş üzerindeki etkileri değerlendirilmiştir. ticari CLEA teknolojisi kullanılarak çapraz bağlanan ticari selülaz formülasyonunun, *P. pastoris*'te üretilen rekombinant EGI ve EGI\_L5 enzimlerinin viskon kumaşların boncuklanma notlarını % 20 oranında iyileştirdiği ve kumaş mukavemetinde kayıplara yol açmadığı bulunmuştur.

## ACKNOWLEDGMENTS

I would specially like to thank my supervisor, Assoc. Prof. Osman Uğur Sezerman for his guidance and moral support throughout the study and for providing me the oppurtunities in Sabancı University. I would like to thank DPT, TUBITAK and Sabancı University Research Funds for financial support of this research.

I would like to thank Prof. Yusuf Menceloğlu and Asst. Prof. Alpay Taralp for their guidance throughout the study and Prof. Canan Atılgan for her input throughout modelling studies, to Prof. Selim Çetiner, Assoc. Prof. Hikmet Budak and Prof. Pervin Anış for their contributions to my thesis as jury members.

I would especially like to thank Dr. Özgür Gül for his support and help in the laboratory, for his contributions to this thesis; to Ms. Emel Durmaz, Mr. Cem Meydan Mr. Aydın Albayrak, Mr. Tuğsan Tezil, Dr. Çağrı Bodur, Dr. Ayça Çeşmelioglu for their help and kind moral support. I would like to thank Mr. Eren Şimşek for SEM analysis. I would specially like to thank to all SezermanLab trainees starting from Begüm Topçuoğlu, Batuhan Orbay Yenilmez, Selcan Tuncay, İnci Ökten, Başak Arslan, Hazal, Duygu, Burcu, Serkan ... who had given their unconditional support and provided help throughout this thesis. I would like to thank to Ms. Ayşe Özlem Aykut, Ms. Aslı Çalık and Mr. Recep Aydın for their contributions to some parts of this work. I would also give my special thanks to Hürmet Turan on behalf of Denge Kimya for the fabric tests.

I am very grateful to my dear mother and father for their trust, support and love, for teaching me the scientific way of thinking, reasoning and criticizing. I am grateful to my brother Gökhan, for his support and for all the things we shared throughout our lives. I am grateful to Özden for her kind support throughout this thesis. I am thankful to Esen, Hüseyin and Yankı for their love and support and for the numerous times they have driven me to the university during weekends. Finally, I am very grateful to Tekin for his love and support, for being always there for me.

## TABLE OF CONTENTS

CHAPTER 1 .....	16
1. BACKGROUND .....	16
1.1. Cellulolytic System of <i>Trichoderma reesei</i> .....	17
1.1.1. Endoglucanases, (EG) (endo-1,4- $\beta$ -glucanase, 1,4- $\beta$ -D-glucan-4-glucanohydrolase, EC 3.2.1.4).....	22
1.1.1.1. Endoglucanase I, Cel7B.....	22
1.1.1.1. Endoglucanase III, Cel12A.....	23
1.1.2. Cellobiohydrolases (CBH) (exo-1,4- $\beta$ -glucanase, 1,4- $\beta$ -D-glucan cellobiohydrolase, EC 3.2.1.91).....	24
1.1.3. $\beta$ -glucosidase .....	25
1.2. Biopolishing.....	27
1.3. Cellulose and Viscose.....	29
1.4. Protein Engineering .....	32
1.4.1. Rational Design of Proteins .....	33
1.4.2. Directed Evolution (Molecular Evolution).....	34
1.5. Crosslinked Enzyme Aggregates (CLEA).....	35
CHAPTER 2 .....	38
2. PURPOSE.....	38
3. MATERIALS AND METHODS.....	39
3.1. Molecular Modeling .....	39
3.1.1. Molecular Models .....	39
3.1.2. Molecular Dynamics (MD) Simulations.....	40
3.1.3. Molecular Mechanics (MM) Simulations.....	40
3.2. Microorganisms, enzymes and chemicals .....	41
3.3. Site directed mutagenesis.....	41
3.4. Cloning.....	43
3.5. Transformation and Screening.....	44
3.6. Copy Number Determination.....	45
3.7. Small Scale Expression of recombinant <i>Pichia pastoris</i> strains .....	45
3.8. Bioreactor Cultivations of recombinant <i>Pichia pastoris</i> strains.....	46

3.9.	SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Zymogram Analysis .....	46
3.10.	Purification of Recombinant Proteins.....	47
3.11.	Protein Assays.....	47
3.12.	Enzyme Assays .....	47
3.12.1.	Effect of Temperature on Enzyme Activity.....	47
3.12.2.	Effect of pH on Enzyme Activity .....	48
3.12.3.	4-MUC Assays.....	48
3.12.4.	Stability assays.....	48
3.13.	CLEA Preparation.....	49
3.13.1.	CLEA preparation from commercial cellulase .....	49
3.13.2.	CLEA preparation from EGI and EGI_L5 .....	49
3.14.	Enzymatic Biofinishing of Viscose Fabrics.....	50
3.15.	Light Microscope and SEM Characterization .....	51
3.16.	Pilling Test.....	51
3.17.	Bursting Strength Test .....	51
4.	RESULTS .....	52
4.1.	Modelling and production of <i>Trichoderma reesei</i> endoglucanase 1 and its mutant in <i>Pichia pastoris</i> .....	52
4.1.1.	Molecular Modelling .....	52
4.1.2.	Production of Recombinant Enzymes and Fermentation .....	57
4.1.1.	Purification of Recombinant Proteins.....	60
4.1.1.	Zymogram Analysis.....	61
4.1.1.	Activity and Stability Analysis .....	61
4.2.	Effect of codon optimization on the production of EGI in <i>Pichia pastoris</i> .....	64
4.2.1.	Production of Recombinant Enzymes and Fermentation .....	64
4.2.2.	Purification of Recombinant Proteins.....	70
4.2.3.	Activity and Zymogram Analysis.....	70
4.3.	Cloning and Production of Recombinant Enzymes and Mutants.....	71
4.3.1.	Production of EGI_BC.....	71
4.3.2.	Production of EGIII .....	76
	Figure 45: pPic $\alpha$ A-egl3 plasmid map (drawn with VectorNTI).....	77
4.3.3.	Production of CBHI.....	79

4.4.	CLEA .....	82
4.5.	Fabric Tests.....	85
5.	DISCUSSION .....	91
5.1.	Modelling and production of <i>Trichoderma reesei</i> endoglucanase 1 and its mutant in <i>Pichia pastoris</i> .....	91
5.2.	Effect of codon optimization on the production of EGI in <i>Pichia pastoris</i> ....	92
5.3.	Cloning and Production of Recombinant Cellulases in <i>P. pastoris</i> .....	94
5.4.	CLEA and Biopolishing.....	95
6.	CONCLUSION.....	97

## LIST OF FIGURES

Figure 1: 1EG1, The crystal structure of the catalytic core domain of endoglucanase I (CEL7B) from <i>Trichoderma reesei</i> at 3.6 Å resolution.....	23
Figure 2: 1H8V, the crystal structure of endoglucanase III (Cel12A) from <i>Trichoderma reesei</i> .....	24
Figure 3: Molecular model showing cellulose hydrolysis by CEL7A.....	25
Figure 4: 1CEL, the crystal structure of the catalytic core of cellobiohydrolase I (Cel7A) from <i>Trichoderma reesei</i> .....	25
Figure 5: Endo-exo synergism between endoglucanases and cellobiohydrolases during cellulose hydrolysis.....	27
Figure 6: Repeating unit of a cellulose molecule .....	31
Figure 7: Cross-section of a viscose fiber (Lenzing Viscose® 2,8 dtex) and its schematic representation.....	32
Figure 8: Crosslinking of a protein with a homobifunctional crosslinker, glutaraldehyde .....	36
Figure 9: Overlap extension PCR schematics for <i>egl1_L5</i> . <i>egl1_L5</i> mutant gene (bottom drawing) prepared from <i>egl1s</i> (top drawing) .....	42
Figure 10: Overlap extension PCR schematics for <i>egl1_BC</i> . <i>egl1_BC</i> mutant gene (bottom drawing) prepared from <i>egl1s</i> (top drawing) .....	43
Figure 11: RMSD (Å) of EGI, EGI_L1, EGI_L2, EGI_L3, EGI_L4, EGI_L5, EGI_L6, EGI_L7 along simulation time (ps) during 4ns MD simulations at 450 °K.....	53
Figure 12: RGYR (Å) of EGI, EGI_L1, EGI_L2, EGI_L3, EGI_L4, EGI_L5, EGI_L6, EGI_L7 along simulation time (ps) during 4ns MD simulations at 450 °K.....	54
Figure 13: RMSD (Å) of EGI vs EGI_L5 along simulation time (ps) during 10 ns MD simulations at 300 °K.....	54
Figure 14: Distance of active site residues to each other during 4 ns MD simulations of EGI and EGI_L5. ....	55
Figure 15: Superimposed structures of EGI and loop mutant EGI_L5. L5 was inserted between residues 112 <sup>th</sup> and 113 <sup>th</sup> of EGI. ....	55
Figure 16: EGI and EGI_L5 stability coefficients calculated for each residue from molecular mechanics simulations .....	56
Figure 17: EGI and EGI_L5 stability coefficients calculated for each residue from molecular mechanics simulations (for simplicity, stability coefficients for residues 100-230 are shown).....	57
Figure 18: Overlap extension PCR results for the production of <i>egl1_L5</i> gene.....	58
Figure 19. Zymogram (upper picture) and SDS-PAGE (lower picture) analysis of fermentation products .....	59
Figure 20: Fermentation data analysis for EGI fed-batch fermentation .....	59
Figure 21: Fermentation data analysis for EGI_L5 fed-batch fermentation.....	59
Figure 22: SDS-PAGE of affinity batch purified EGI and EGI_L5.....	60
Figure 23: Activity of purified EGI and EGI_L5 at the same protein concentration against 4-MUC at 45 °C.....	60
Figure 24: Effect of temperature on hydrolysis of 0.5 % CMC (w/v) by EGI and EGI_L5 .....	62
Figure 25: Effect of pH on hydrolysis of 0.5 % CMC (w/v) by EGI and EGI_L5 .....	62
Figure 26: Residual activity of EGI and EGI_L5 at 50 °C upon incubation for 0 to 72 hours at 50 °C against CMC at pH 4.8 .....	63
Figure 27: Residual activity of EGI and EGI_L5 at 50 °C upon incubation for 0 to 2 hours at 70 °C against CMC at pH 4.8. ....	63

Figure 28: Colony PCR for subcloned <i>eglIs</i> (lanes 2-4) and <i>eglI</i> (lanes 6 and 7) genes using 5' and 3' AOX primers .....	66
Figure 29: Colony PCR for <i>eglIs</i> and <i>eglI</i> expression cassettes transformed into <i>Pichia pastoris</i> KM71H using 5' and 3' AOX primers .....	66
Figure 30: Zymogram analysis of expressed EGIs and EGI in shake flasks against 4-MUC.. .....	67
Figure 31: Zymogram analysis of expressed EGIs and EGI as batch fermentation products against 4-MUC .....	67
Figure 32: Activities of EGIs and EGI against 4-MUC throughout batch fermentation.. .....	68
Figure 33: Activities of EGIs and EGI against 4-MUC throughout fed batch fermentation. ....	69
Figure 34 : SDS-PAGE of affinity batch purified EGIs and EGI.....	69
Figure 35: a) Effect of temperature on recombinant EGI activity against CMC b) Effect of pH on recombinant EGI activity against CMC .....	71
Figure 36: Overlap extension PCR results for the production of <i>eglI_bc</i> gene .....	71
Figure 37: Colony PCR for <i>eglI_bc</i> expression cassette transformed into <i>Pichia pastoris</i> KM71H using 5' and 3' AOX primers .....	72
Figure 38: Fermentation data analysis for EGI_BC fed-batch fermentation (pH 5, 29 °C) .....	73
Figure 39: Fermentation data analysis for EGI_BC fed-batch fermentation (pH 7, 29 °C) .....	73
Figure 40: Fermentation data analysis for EGI_BC fed-batch fermentation (pH 5, 25 °C) .....	74
Figure 41: SDS-PAGE and zymogram analysis of expressed EGI_BC as fed-batch fermentation products against 4-MUC at pH 5, 29 °C.....	74
Figure 42: Zymogram analysis of expressed EGI_BC as fed-batch fermentation products against 4-MUC at pH 7, 29 °C (top picture), at pH 5, 25 °C (bottom picture) .....	74
Figure 43 : SDS-PAGE of affinity batch purified EGI_BC .....	75
Figure 44: a) Effect of temperature on recombinant EGI_BC activity against CMCb) Effect of pH on recombinant EGI_BC activity against CMC .....	75
Figure 45: p <i>PiczαA-egl3</i> plasmid map .....	77
Figure 46: Colony PCR amplification of <i>pPiczαA_egl3</i> harboring <i>P. pastoris</i> (KM71H) clones with AOX primers .....	77
Figure 47: Azo-CMC activity of different <i>Pichia pastoris</i> clones producing EGIII.....	77
Figure 48: Fermentation data analysis for EGIII clone C13 fed-batch fermentation .....	78
Figure 49: SDS-PAGE and Zymogram analysis of expressed EGIII as fed-batch fermentation products against 4-MUC .....	78
Figure 50: Effect of temperature on recombinant EGI_BC activity against CMC .....	79
Figure 51: Colony PCR amplification of <i>pPiczαB_cbh1</i> harboring <i>P. pastoris</i> (KM71H) clones with AOX primers .....	80
Figure 52: a) Fermentation data analysis for CBHI clone Y2 fed-batch fermentation. b) SDS-PAGE analysis of expressed CBHI as fed-batch fermentation products against 4-MUC .....	80
Figure 53: a) Effect of temperature on recombinant CBHI activity against CMC b) Effect of pH on recombinant CBHI activity against CMC.....	81
Figure 54: Effect of pH on Gempil 4L activity after precipitation. ....	82
Figure 55: Effect of glutaraldehyde concentration on CLEA activity against CMC after crosslinking.....	83
Figure 56: Effect of CLEA size on CLEA activity.....	83



Figure 57: Effect of glutaraldehyde concentration on CLEA prepared from EGI and EGI_L5. ....	84
Figure 58: Effect of crosslinking on enzyme activity at pH 5, 55°C. ....	84
Figure 59: Digital light microscope photographs of viscose knitted fabrics under ~15X and ~400X magnification. ....	87
Figure 60: Digital light microscope photographs of viscose knitted fabrics teated with 1X and 4X CLEA under ~15X and ~400X magnification. ....	87

## LIST OF TABLES

Table 1: <i>Trichoderma reesei</i> cellulolytic system components .....	20
Table 2: Overlap PCR extension primers designed for <i>eglI_L5</i> gene.....	42
Table 3: Overlap PCR extension primers designed for <i>eglI_BC</i> gene.....	43
Table 4: Plasmid names and linearization sites for all genes. ....	45
Table 5: Kinetic constants for EGI and EGI_L5 calculated from their activity against 4-MUC at 45 °C. ....	64
Table 6: Comparison of codon usage in <i>Pichia pastoris</i> genes with that in native and codon-optimized <i>eglI</i> genes. ....	65
Table 7: Pilling test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) or recombinant enzymes.....	86
Table 8: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different treatment times or recombinant enzymes .....	86
Table 9: Pilling test results for viscose woven fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C).....	88
Table 10: Bursting strength test results for viscose woven fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C). ....	88
Table 11: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different sizes in hybridization chamber.....	89
Table 12: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different sizes in special apparatus .....	90

## ABBREVIATIONS

AATCC	Association for American Textile Chemists and Colorists
BSA	Bovine serum albumin
C $\alpha$	Carbon alpha
CBD	Cellulose binding domain
CBM	Carbohydrate binding module
CBH	Cellobiohydrolase
CDW	Cell dry weight
CLEA	Crosslinked enzyme aggregates
CLEC	Crosslinked enzyme crystals
CMC	Carboxymethyl cellulose
DNS	Dinitrosalicylic acid
EG	Endoglucanase
GH	Glycosyl hydrolase
K <sub>m</sub>	Substrate concentration where half maximal velocity of an enzymatic reaction is reached
LR	Liquor ratio
MD	Molecular dynamics
MM	Molecular mechanics
4-MUC	4-Methylumbelliferyl beta-D-cellobioside
4-MUL	4-Methylumbelliferyl beta-D-lactoside
NaOAc	Sodium acetate
RGYR	Radius of gyration
RMSD	Root-mean-square deviation
RPM	Revolutions per minute
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	Scanning Electron Microscopy
V <sub>max</sub>	Maximum velocity of the enzymatic reaction

## CHAPTER 1

### 1. BACKGROUND

Enzymes are used extensively in the industrial processes along with conventional chemical processes. As a result of green technology, extensive research has been done to replace conventional chemical processes with environmentally friendly, less harmful alternatives. Enzymes are used for a wide range of industrial applications that include biofuels, detergent, paper, food, feed, pharmaceutical and textile industries. As new industrial application areas emerge, there is an increasing demand for the production of enzymes that can be used for specific purposes and that can withstand heavy chemical conditions and elevated temperatures of the industrial processes. One of these rapidly growing areas is the textile processing industry. Enzymes are preferred in textile processes because of their specificity, speed, biodegradability, operational stability and vast application areas. They are especially used in biofinishing and biopreparation of the textiles. Cellulases are commonly used in biofinishing of cellulosic fabrics. They remove the microfibrils on the fabric surface and prevent formation of the pills. They are routinely used for the removal of pills from cotton and viscose fabrics. They are effective in removal of pills from cotton fabrics but not from viscose fabrics. Moreover, they cause a reduction in the tensile strength of the fabrics due to their aggressive action. Our previous studies have shown that application of crosslinked commercial cellulases had prevented the loss of tensile strength in viscose fabrics but the crosslinked commercial cellulases did not prevent the pilling formation (Bayram Akcapinar, 2005). Moreover, it has been found that the activity of the commercial enzyme was lowered after crosslinking. Cellulases are multi-component enzymes. Their activity on fabrics and their effects on fabric surfaces change according to cellulase composition. Therefore, there is still need for production of mono-component cellulases and cellulase formulations that can improve viscose fabric pilling properties without a

reduction in fabric strength, enzyme stability and activity. This can be achieved by producing cellulase components in homologous or heterologous hosts; increasing the enzyme size either by protein engineering or by chemical modification or a combination of both. To this end,

- EGI, EGIII and CBHI of *T. reesei* were cloned and expressed in *P. pastoris*.
- A loop mutant of EGI was prepared by introduction of a ten aminoacid long loop by molecular modelling and site directed mutagenesis for the creation of hotspots for directed crosslinking of the enzyme and this enzyme was crosslinked using CLEA method.
- Effect of codon optimization on EGI production was analyzed.
- A mutant of EGI was prepared by inserting a second catalytic domain to EGI and thereby forming a bicatalytic mutant of EGI (EGI\_BC) with increased molecular weight.
- All of the recombinant enzymes were produced in a laboratory scale fermenter and characterized.
- A commercial cellulase preparation was crosslinked using crosslinked enzyme aggregate (CLEA) technology and fractionated according to the particle size.
- Effects of native, engineered and chemically modified cellulases on viscose fabrics were evaluated.

### **1.1. Cellulolytic System of *Trichoderma reesei***

The very first strain of fungi that is capable of hydrolyzing cellulose was first discovered during World War II when it was noticed that the rate of deterioration of cellulosic materials that belong to the U.S. Army was increased in the South Pacific. In order to produce a solution to this problem immediate research was started and as a result the first strain QM6a was isolated. This strain was first identified as *Trichoderma viride* and later recognized as *Trichoderma reesei* (Bhat, 2000). The research on this cellulose degrading organism and cellulose degradation has started. In the last half of the 20<sup>th</sup> century there has been a remarkable progress in isolation of microorganisms producing cellulases; improving the yield of cellulases by mutation; purifying and characterizing the cellulase components; understanding the mechanism of cellulose

degradation cloning and expression of cellulase genes; determining the 3-D structures of cellulase components; understanding structure-function relationships in cellulases; and demonstrating the industrial potential of cellulases (Kumar, Yoon, & Purtell, 1997).

Cellulose, being the most abundant polymer on earth, has a very important place in bioethanol production and it is a starting material for many fabrics used in textile (Zhang & Lynd, 2004). Due to these properties of cellulose, cellulases, enzymes that degrade cellulose, are becoming very important materials for biotechnology and enzyme engineering (Bhat & Bhat, 1997). Among all of the enzymes cellulases are being used increasingly for a variety of industrial purposes (Bhat, 2000) and consequently, a lot of effort has been put into their cloning (Qin, Wei, Liu, Wang, & Qu, 2008; Saloheimo, Nakari-Setälä, Tenkanen, & Penttilä, 1997; Yao, Sun, Liu, & Chen, 2008) and expression (Aho, 1991) as well as their study by site-directed mutagenesis (Nakazawa, et al., 2009; Rignall, et al., 2002; Sandgren, Stahlberg, & Mitchinson, 2005). Cellulases are multicomponent enzymes. There are three major types of cellulases secreted by *Trichoderma reesei*: Endoglucanases, 1,4- $\beta$ -D-glucan 4-glucanohydrolases; Cellobiohydrolases, 1,4- $\beta$ -D-glucan cellobiohydrolases; Cellobiases,  $\beta$ -D-glucosidases. *Trichoderma reesei* has at least six endoglucanases, two cellobiohydrolases, and two  $\beta$ -D-glucosidases (Bhat & Bhat, 1997; Heikinheimo, 2002). EGI and EGII are the main components of the *T. reesei* endoglucanases and they comprise ~10 % of the secreted proteins of the organism (Heikinheimo & Buchert, 2001).

Application of cellulases in industrial processes has increased to a considerable amount in the last thirty years. Cellulases are used in textile industry for biopolishing of textiles and fabrics to improve fabric quality (Videb, 2000), biostoning of denim garments to obtain a fashionable aged appearance (Cavaco-Paulo, Almeida, & Bishop, 1996). They are also used extensively in feed, food industries, in pulp and paper processing, in laundry (Bhat, 2000). In the last decade, several studies focused on the use of cellulases for the conversion of lignocellulosic biomass to produce biofuels as an alternative renewable energy source to fossil fuels (Bayer, Lamed, & Himmel, 2007; Percival Zhang, Himmel, & Mielenz, 2006; Wilson, 2009).

Cellulases are enzymes that work under extreme conditions of textile industry: e.g. elevated temperatures and pH, exposure to organic solvents and various chemicals. Engineering of the cellulase enzyme should not cause a dramatic reduction in the

enzyme activity and stability since enzymes stable under such harsh conditions attract attention due to their applicability in industrial processes. Understanding the underlying mechanisms of stability may help designing an engineered cellulase without a reduction in the enzyme stability and activity.

Cellulolytic enzymes are produced by a wide variety of organisms. Few of these enzymes are capable of degrading crystalline cellulose effectively. Among these microorganisms, the extremophilic ones are very important because of the stability of their enzymes under harsh conditions such as highly acidic and alkaline pHs as well as temperatures up to 90 °C (Lamed & Bayer, 1988). Important thermophilic microorganisms capable of degrading cellulose are *Clostridium thermocellum*, *Thermomonospora fusca*, *Thermoascus aurantiacus*, *Sporotrichum thermophile*, *Humicola insolens* and *Chaetomium thermophile*. *Clostridium thermocellum* differs from other cellulolytic microorganisms since it secretes all its cellulolytic enzymes in a protein complex called cellulosomes (Németh, et al., 2002). Most extensive research about the cellulases has been done on aerobic fungi such as *Trichoderma koningii* (Halliwell & Vincent, 1981) and *T. reesei* (Liming & Xueliang, 2004; Medve, Karlsson, Lee, & Tjerneld, 1998; Miettinen-Oinonen, Paloheimo, Lantto, & Suominen, 2005).

Cellulases are the only enzymes used in biofinishing of the cotton fabrics. These enzymes are suitable for wet processes and they can be used almost in all textile machines. Nowadays commercial cellulase preparations for different types of fabrics are available for use in biopolishing. They exhibit a wide range of pH and temperature stability and activity. Commercial cellulase preparations are mostly from the filamentous fungi, *Trichoderma reesei*. Cellulases are extracellular enzymes. They are secreted out of the cells. Industrial producers take this advantage into consideration. That is why *Trichoderma reesei* is the workhorse of industry in terms of production of cellulases. *Trichoderma reesei* produces cellulases in large quantities and secretion of the enzyme components allows rapid purification of the enzymes.

Researchers are also working on producing genetically modified cellulase enzymes with the desired properties for different types of processes. Directed evolution and site directed mutagenesis studies which target the cellulases are reported (Sandgren, et al., 2005). Moreover, site-directed mutagenesis studies have been used for the

identification of active site residues and residues responsible for the stability of the cellulases. These studies are very valuable tools because they provide the information for the design of new cellulases having specific activities.

Cellulases are multicomponent enzymes. There are three major types of cellulases secreted by *Trichoderma reesei*: Endoglucanases, 1,4- $\beta$ -D-glucan 4-glucanohydrolases; Cellobiohydrolases, 1,4- $\beta$ -D-glucan cellobiohydrolases; Cellobiases,  $\beta$ -D-glucosidases [48]. *Trichoderma reesei* has at least six endoglucanases, two cellobiohydrolases, and two  $\beta$ -D-glucosidases (Bhat & Bhat, 1997; Heikinheimo & Buchert, 2001). Figure 1.12 indicates the molecular weights and number of aminoacids of some of the cellulase components.

Table 1: *Trichoderma reesei* cellulolytic system components (Vinzant, et al., 2001).

Name acc. to GH Classification*	Cellulase components of <i>Trichoderma reesei</i>	Molecular Weight (kDa)	Number of amino acids	Amount of total cellulase**
<b>CEL7B</b>	<b>EG I</b>	48,2	459	9
<b>CEL5A</b>	<b>EG II</b>	44,2	418	8
<b>CEL12A</b>	<b>EG III</b>	23,5	218	<1
<b>CEL45A</b>	<b>EG IV</b>	35,5	344	not known
<b>CEL61</b>	<b>EG V</b>	24,5	242	not known
<b>CEL7A</b>	<b>CBH I</b>	54	513	55
<b>CEL6A</b>	<b>CBH II</b>	49,6	471	18
<b>GH Family 3</b>	<b>B-D-glucosidase I</b>	78,5	744	

\* The names of the cellulases are based on the GH nomenclature system introduced by Henrissat et al. (Henrissat & Bairoch, 1993). \*\*Amount of total cellulase secreted by *T. reesei*.

Cellulases belong to the glycosyl hydrolase family of enzymes. This enzyme family contains nearly 96 subfamilies. Cellulases are present in at least 11 of these subfamilies; GH 5, 6, 7, 8, 9, 12, 26, 44, 45, 48 and 61 according to [www.cazy.org](http://www.cazy.org) (Henrissat, 1991; Henrissat & Bairoch, 1993). Subfamily classifications of the glycosyl hydrolase family are done on the basis of their amino acid sequences. The three dimensional structures and enzyme-substrate interaction mechanisms display some differences in every subfamily.

These multicomponents act synergistically for the degradation of cellulose. They act specifically on 1,4- $\beta$ -glycosidic bonds of the cellulose. Cellulase has two domains. One of them is the catalytic core domain and the other is the cellulose binding domain.



The hydrolysis of the cellulosic substrates takes place inside the catalytic core domain and this domain occupies the largest part of the enzyme. These two domains linked by a short linker peptide form the intact bimodular enzyme (Kleywegt, et al., 1997; Sandgren, et al., 2005). The length of the linker varies from 20 to 40 aminoacids. The linker peptide is rich in Proline, Threonine and Serine residues and it is often O-glycosylated. O-glycosylation provides maintenance of the extended conformation of the linker peptide and also protects the linker region against proteases of the organism. Cellulose binding domains are thought to adsorb to the cellulose thereby acting as anchor points for the whole enzyme. They keep the cellulosic substrate in the vicinity of the enzyme. These adsorption properties of CBD enable the enzyme to have a higher turnover rate (Ståhlberg, 1991).

Cellulose binding domains of fungi, alga and bacteria are classified into two families. The shorter CBDs (30-40 amino acids) are the ones from the fungi are classified as family II and the longer ones (100-150 amino acids) are the ones from bacteria and algae and classified as family I (Reinikainen, Teleman, & Teeri, 1995). They are thought to have arisen by a convergent evolution since they do not have much sequence similarities. But they have conserved aminoacids having aromatic side chains and these are thought to be involved in cellulose binding in all types of CBDs. CBDs are also classified into more than 30 families according to CAZY database ([www.CAZY.org](http://www.CAZY.org)). Much effort has been put to clarify the mechanism of adsorption and its effect on the activity of the cellulase components. Deletion mutants of cellulase components (CBD deleted) were prepared and analysis of their adsorption trends revealed a decrease of 50-80 % of activity of fungal cellulases on insoluble substrates (Srisodsuk, Reinikainen, Penttila, & Teeri, 1993).

It was suggested that in CBHI of *T. reesei* both core domain and CBD participated in binding and in bacteria only CBDs are involved in binding. Site directed mutagenesis directed towards CBDs of CBHI (Y492A, Y492H and P477R) indicated that conserved aromatic aminoacids are essential in binding and it is known that hydrophobic interactions are also important for binding (Reinikainen, et al., 1995). The CBD is located at the C-terminus of the CEL7A, CEL7B, CEL45A, and CEL61A catalytic core domains and at the N-terminus of the CEL5A, and CEL6A catalytic core domains.

CEL12A is the only *T. reesei* cellulase that is known to have no cellulose binding domains.

### **1.1.1. Endoglucanases, (EG) (endo-1,4- $\beta$ -glucanase, 1,4- $\beta$ -D-glucan-4-glucanohydrolase, EC 3.2.1.4)**

Endoglucanases are the endocellulases which randomly hydrolyze the cellulose chains internally. Action of endoglucanases produces new chain ends and changes the degree of polymerization of the cellulose. The target of the endoglucanases is the amorphous cellulose (Heikinheimo & Buchert, 2001). They exhibit lower activity towards insoluble substrates such as crystalline cellulose. There are at least 6 identified endoglucanases in *Trichoderma reesei* (EG I-VI). CEL74A (EG VI) was described only at the protein level (Bower, et al., 1998). CEL7B and CEL5A are the main components of the *T. reesei* endoglucanases and they comprise ~10 % of the secreted proteins of the organism (Heikinheimo & Buchert, 2001). CEL7B, CEL5A and CEL12A cleave  $\beta$ -1,4-glycosidic bonds with retention of anomeric configuration, yielding the  $\beta$ -anomer as the reaction product and CEL45A uses the inverting mechanism. The exact mechanism of CEL61A and EGVI is not exactly known according to the current knowledge. Azavedo et al.(2000), suggested that agitation levels has a profound effect on endoglucanase activity and at high levels of agitation the presence of CBDs is not key to the functioning of the endoglucanases (Azevedo, Bishop, & Cavaco-Paulo, 2000).

#### **1.1.1.1. Endoglucanase I, Cel7B**

CEL7B belongs to the glycosyl hydrolase(GH) family 7. CEL7B is homologous to CEL7A, cellobiohydrolase I with ~45 % sequence identity. The active site of CEL7B is located in an open cleft not crowded by extended loops as in CEL7A. The catalytic residues of the active site are identified as E196-D198-E201-H212 (Kleywegt, et al., 1997). Comparison of the structures of the *T. reesei* EG I and *H. insolens* EG I, reveals that they have similar substrate-binding grooves: both proteins have their active site located in an open cleft (Sandgren, et al., 2005). Figure 1 shows the three dimensional structure of the catalytic core domain of endoglucanase I of *T. reesei*. Estimated molecular weight of CEL7B is 48 kDa and it is known to have a pI around 4.5. Several studies indicate a heterogenous glycosylation of *Trichoderma reesei* EGI (Eriksson, et

al., 2004; Hui, White, & Thibault, 2002). Thus the molecular weight changes according to the glycosylation pattern. EGI is known to have N-linked glycosylation.

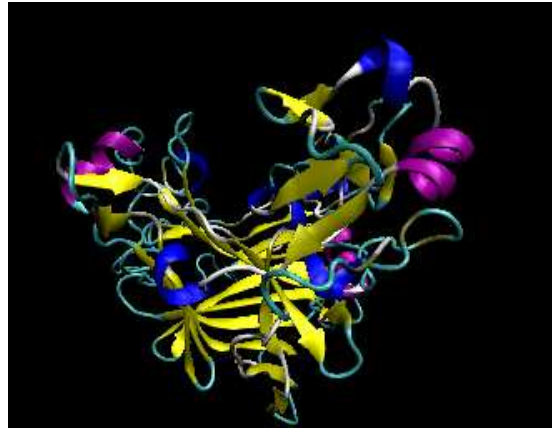


Figure 1: 1EG1, The crystal structure of the catalytic core domain of endoglucanase I (CEL7B) from *Trichoderma reesei* at 3.6 Å resolution (Kleywegt, et al., 1997).

There are a few studies that indicate heterologous production of Cel7B in different hosts. Korhola et al. have introduced *egl1* gene of *T. reesei* into *S. cerevisiae* under the control of *PGK1*, *MEL* and *ADHI* promoters and used the enzyme as a reporter for screening mutagenized yeast strains (Aho, Arffman, & Korhola, 1996). CEL7B cDNA under the control of *ADHI* promoter has produced  $10^{-4}$ - $10^{-5}$  g/l CEL7B. With the screening of first generation and second generation mutants, they were able to isolate mutants producing immunoreactive CEL7B as high as 0,04 g/l but only the 2% of the enzyme was found to be active. In a more recent study (Nakazawa, et al., 2008), CEL7B catalytic core domain was expressed in *E. coli* strains Rosetta-gami B (DE3) pLacI or Origami B (DE3) pLacI. CEL7B produced as a functional intracellular protein in these *E. coli* cells. Maximum productivity for CEL7B catalytic domain was found to be at 15°C with a yield of 6.9 mg/l. CEL7B was found to have a pH optimum around 5. Temperature stability experiments have indicated that the recombinant enzyme has lost all of its activity upon incubation at 60 °C and 70 °C for 15 minutes.

#### 1.1.1.1. Endoglucanase III, Cel12A

Cel12A is a GH family 12 endoglucanase. CEL12A is known to have a molecular weight around 25 kDa and a pI of 7.5 (Hakansson, Fagerstam, Pettersson, & Andersson, 1978; Ulker & Sprey, 1990). The pH optimum for CEL12A was found to be 5.8 and it was shown to exhibit its optimal temperature at around 52 °C (Ulker & Sprey, 1990).

CEL12A is known to have no CBDs and it is known to be sparsely glycosylated. The active site residues were found to be E116 and E200 (Okada, Mori, Tada, Nogawa, & Morikawa, 2000). The crystal structure of CEL12A is shown in Figure 2. CEL12A is produced in very little quantities by *T. reesei* (less than 1%). Nakazawa et al. expressed *egl3* cDNA in *E. coli* (Nakazawa, et al., 2008) and performed directed evolution experiments using error prone PCR on the recombinant enzyme (Nakazawa, et al., 2009). Stability and the specific activity of the enzyme was found to be enhanced by directed evolution. The pH stability experiments of the wild type CEL12A and mutant CEL12A have indicated that directed evolution has broadened the pH range of the enzyme.

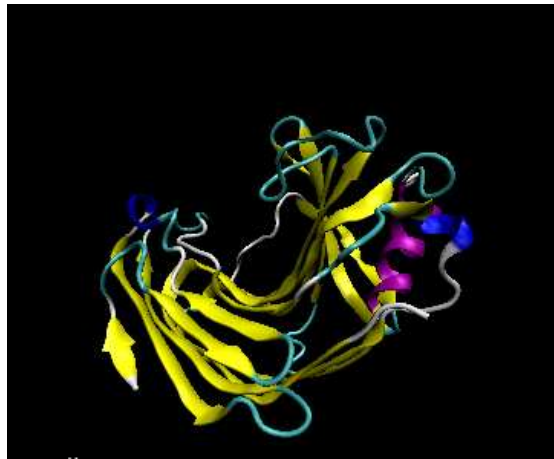


Figure 2: 1H8V, the crystal structure of endoglucanase III (Cel12A) from *Trichoderma reesei* (prepared with VMD).

### 1.1.2. Cellobiohydrolases (CBH) (exo-1,4- $\beta$ -glucanase, 1,4- $\beta$ -D-glucan cellobiohydrolase, EC 3.2.1.91)

Cellobiohydrolases are exocellulases that hydrolyze the cellulose chains from the ends releasing cellobiose as the end product. *T. reesei* has two CBHs. CBH I splits cellobiose from the reducing end and CBH II from the nonreducing end (Heikinheimo & Buchert, 2001).

Structural studies revealed that CBH I core domain contains a 40 Å long tunnel shaped active site along the enzyme molecule. A model for cellulose hydrolysis by CBHI was shown in Figure 3. The tunnel shaped active site explains the high affinity of CBHs to crystalline cellulose during the progressive catalytic cycles. This also explains the processivity seen in CBHs. The loops present on the surface of CBH allows the

extricated cellulose chain from adhering back to the crystalline cellulose. Moreover, since the crystalline cellulose is the highly ordered one, it can fit easily into that tunnel whereas amorphous cellulose having a more loose structure can not easily fit to the same cavity. Figure 4 indicates the three dimensional structure of CBH I from *T. reesei*. CBH I has a retaining mechanism of hydrolysis whereas CBH II has inverting mechanism (Sandgren, et al., 2005).

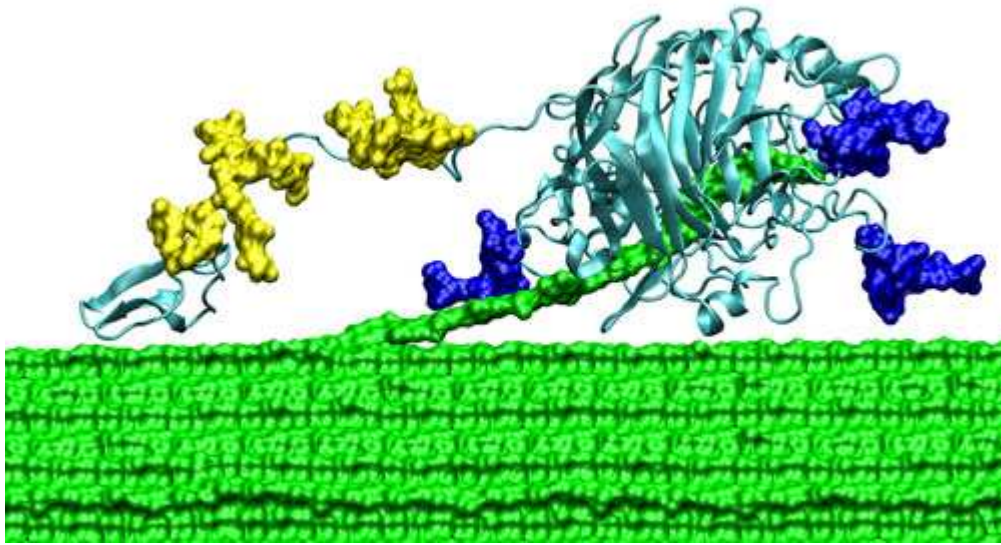


Figure 3: Molecular model showing cellulose hydrolysis by CEL7A. N-linked glycosylation is shown in blue color and O-linked glycosylation is indicated with yellow color. Cellulose molecule is shown in green. (<http://www.nrel.gov/biomass/images/cbh1.jpg>)



Figure 4: 1CEL, the crystal structure of the catalytic core of cellobiohydrolase I (Cel7A) from *Trichoderma reesei* (prepared with VMD) (Divne, et al., 1994).

### 1.1.3. $\beta$ -glucosidase

$\beta$ -glucosidases also called cellobiases are responsible for the hydrolysis of cellobiose produced by CBH enzymes into glucose. The function of  $\beta$ -glucosidase is very important. It is known that CBH and EG are inhibited by cellobiose (Gruno, Valjamae, Pettersson, & Johansson, 2004). So the main function of the  $\beta$ -glucosidase in the cellulase system is to overcome the product inhibition of CBH and EGs (Lenting & Warmoeskerken, 2001). The three dimensional structure of *T. reesei*  $\beta$ -glucosidase 1 has not been solved yet.

There exists a strong synergism between cellulase components. The types of synergisms reported upto now are endo-endo, endo-exo, exo-exo, endo-exo-glucosidase exo/endo-glucosidase. However, endo-exo synergism is the most extensively studied one. There is also an intramolecular synergy between the CBD and core catalytic domains. The degree of synergism is different for each type of substrate. For example, endo-exo synergism is mostly pronounced for the degradation of the crystalline cellulose. Degree of synergism was found to be most for cotton and then Avicel and least for acid swollen amorphous cellulose (Zhang & Lynd, 2004). Figure 5 shows the synergistic action of cellulases on cellulose. In this synergism, endoglucanases adsorb to the cellulose microfibrils and start to make internal cuts in the cellulose chain. Then cellobiohydrolases start the hydrolysis from the newly created chain ends and CBHs hydrolyze the cellulose chain processively. This combined action increases the efficiency and activity of the whole cellulase system with respect to the enzymes acting alone.

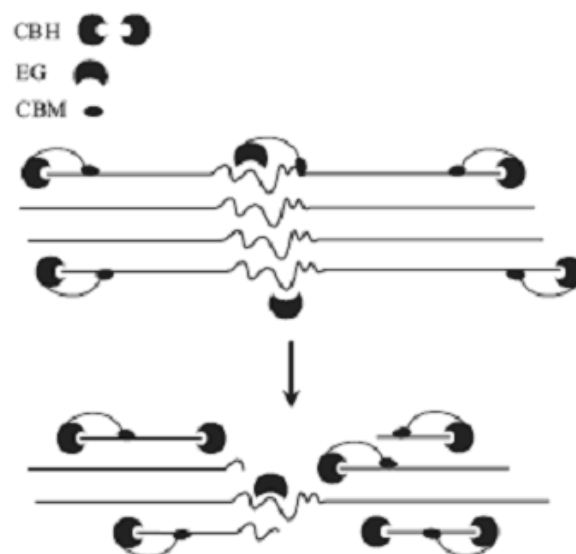


Figure 5: Endo-exo synergism between endoglucanases and cellobiohydrolases during cellulose hydrolysis (Heikinheimo, 2002).

## **1.2. Biopolishing**

Biopolishing also known as biofinishing refers to the removal of the cellulose fibrils and microfibrils protruding from the surface of the fabric or fibers by the action of cellulases. These fibrils and microfibrils are termed as fuzz. These loose microfibrils and fibrils tend to agglomerate on the fabric surface. These loose agglomerations are called pills. Pills are formed during fabric processing in the production plant, washing and/or wearing. The mechanical action provided by the friction of the fabrics during wearing causes pill formation. There is an increasing demand on the use of cellulases in the textile industry for the removal of pills and fuzz formed on the surface of the fabric. Enzymatic treatment provides fabrics with

- better surface properties and look
- improved hand properties
- improved drapeability (ability to hang or stretch out loosely)
- increased brightness
- reduced pilling and pilling tendency
- increased softness compared to the conventional softeners

Biopolishing involves the enzymatic treatment of the cellulosic fabrics such as cotton, linen, rayon and Lenzing's Lyocell and viscose with cellulases that eventually leads to the weakening of the fibers protruding from the surface of the fabric and the removal of the weakened fibers with mechanical action. The tendency for the formation of pills during wearing and washing is minimized since the protruding fibrils are removed by biopolishing. The biopolishing process was patented in 1993 by Videbaek and Andersen (Videbaek & Andersen, 1993) and it is mainly designed to improve fabric quality.

Since biopolishing is an enzymatic process it can be carried out during the wet processing stages. It is mostly performed after bleaching before dyeing. After bleaching, the fabric becomes cleaner and more hydrophilic. So it becomes more prone to attack by cellulases. Biopolishing is not performed after dyeing since there is risk of color fading

and the chemical content of the dyes can reduce the performance of the enzymes by interfering with them. Direct and reactive dyes have been known to have an inhibitory effect on cellulases.

Biopolishing is mostly performed in machines such as a jet-dryer or winch. Enzyme dosage is a very important parameter for having the desired effect. The dosage was determined as a percentage of the garment weight. Usually, 0.5-6 % enzyme over fabric weight is used by the manufacturers. Process parameters such as pH, temperature and duration is determined according to the properties of the cellulase enzyme to be used. Generally the process is performed at pH 4.5-5.5, temperature between 40-55 °C for 30-60 minutes and the enzyme is inactivated usually by increasing the temperature above 80 °C or pH above 10. Soda ash is used for the pH adjustments.

Many aspects of cotton biopolishing with cellulases were studied. For example, Miettinen-Oinonen et al. developed different cellulase formulations (CBH I, CBH I and II, EGII, EGI and II enriched and wild type) by genetic engineering and applied these on the biofinishing of cotton fabrics (Miettinen-Oinonen, et al., 2005). They found that EGII enriched and EG enriched cellulase formulations improved the surface appearance more than CBHI, CBHII and CBH enriched cellulase formulations. All the pilling values were better than the wild type and CBHII was found to be the most effective throughout all CBHs. The pilling values for EGII enriched cellulase formulation was 4.3 and for CBHI enriched cellulase and wild-type cellulase pilling values were 2.3 where a pilling value of 5 indicates no pills and 1 indicates intense pilling.

Although there are many studies on the biopolishing of cotton fabrics, there are a few on the biopolishing of the regenerated cellulose fabrics such as Lyocell and viscose. Use of cellulases in biopolishing of viscose was studied by Ciechańska et al. (Ciechańska, Struszczyk, Miettinen-Oinonen, & Strobin, 2002). Different formulations of cellulases (EGII, CBHI and total cellulase enriched with EGII) and a commercial cellulase (Econase CE, Rohm Enzymes Inc.) from *T. reesei* were applied on two types of viscose woven fabrics. The microscopic properties of the fabrics and residual fibers were analyzed, but the pilling values or pilling tendencies were not evaluated. It was found that use of the commercial enzyme removed most of the microfibrils and fuzz



protruding from the surface of the viscose woven fabric whereas the purified components did not improve the surface of the viscose woven fabric.

According to our knowledge, there are no studies on the biopolishing of knitted viscose fabrics and their pilling values upon enzymatic treatment. Liu et al. analyzed the effects of different commercial and experimental cellulase preparations (commercial multicomponent acid cellulase and monocomponent acidic endoglucanase, experimental EG enriched cellulase) on cotton interlock (type of a stretchable fabric) knitted fabric (J. Liu, et al., 2000) (cited in (Heikinheimo, 2002)). According to their results and interpretation, cellulases were found to have different selectivities when their ratios of pilling to bursting strength, their sensitivities to liquor ratio and mechanical agitation created by the equipment, fiber types were considered. Kumar et al. suggested that EG enriched cellulases had some advantageous properties such as improved hand compared to total cellulase preparations (Kumar, et al., 1997).

One of the problems encountered during biopolishing is the loss of fiber or fabric strength as a result of the aggressive action of the enzymes. This problem is predominantly seen in the biofinishing of lyocell and viscose fabrics. These problems are solved using different formulations of cellulases (Kumar, et al., 1997). Viscose fabrics' tensile strength is known to be lowered when it is wet. This poses an important problem since most of the textile processes are wet processes. Moreover, aggressive cellulases are used in most of the processes. There are commercial cellulase preparations suitable for lyocell biofinishing and most of them are also suggested for biofinishing of viscose fabrics. But to our knowledge there are no specific commercial cellulases for viscose and the ones that are used for lyocell, cotton are insufficient for the removal of the pills on viscose (especially, viscose knitted fabric).

### **1.3. Cellulose and Viscose**

Cellulose is an unbranched polymer of  $\beta$ -1,4-linked glucose molecules. Plants are the only producers of cellulose. It is the most abundant polymer on earth. The glucose units forming the cellulose chain are in six membered pyranose ring. There forms an acetal linkage between the C1 of one pyranose ring and C4 of the next pyranose ring. A single oxygen atom joins the two pyranose rings. In the formation of an acetal, one

molecule of water is lost when an alcohol and a hemiacetal reacts. That is why the glucose units forming the cellulose polymer are also called anhydroglucose units.

The spatial arrangement of the acetal linkages are very important in determining the characteristics of the cellulose molecule. With the formation of pyranose ring, there exists two possibilities for the configuration. The hydroxyl group present on C4 can approach the C1 hydroxyl group from both sides. This results in two different stereochemistries. If the C1 hydroxyl group is on the same side with C6 hydroxyl group, the configuration is called the  $\alpha$ , if they are on the opposite sides the configuration is called the  $\beta$ . Cellulose is known to be in  $\beta$  configuration and is a poly[ $\beta$ -1,4-D-anhydroglucopyranose]. In  $\beta$  configuration all the functional groups (hydroxyls) are in equatorial positions which means they protrude laterally from the extended molecule. These protruding hydroxyl groups are readily available for hydrogen bonding (interchain and intrachain hydrogen bonding are observed). Moreover, inter and intrachain H-bonding and Van der Waals interactions force the cellulose chains into a parallel alignment and finally to an ordered crystalline structure. This property allows the chain of cellulose to extend in a straight line and makes cellulose a good fiber forming polymer by giving tensile strength along the fiber axis. Hydrogen bonding causes the formation of highly ordered crystal structure. This highly ordered crystalline regions are thought to be intruded with less ordered amorphous regions (Zhang & Lynd, 2004). The interchain hydrogen bonds in the crystalline regions gives the fibers their strength and insolubility. In the less ordered regions the cellulose chains are more loose and further apart as a result. This enables the hydroxyl groups to form hydrogen bonds for example with water molecules and causes these regions to absorb water. On the other hand, amylose has the  $\alpha$  configuration, C1 oxygens are in  $\alpha$  configuration. This causes the formation of the linkages between the adjacent glucopyranose residues to be in axial positions and this forces the amylose chain to assume a helical structure maintained by interchain hydrogen bonds. Since helical structure is not proper for fiber formation, starch is not a suitable fiber-forming molecule.

Figure 6 indicates the repeating unit of a cellulose molecule. The repeating unit of cellulose is the anhydrocellobiose. Cellobiose is formed from two identical but 180° rotated anhydroglucose units. This introduces the symmetry to the cellulose molecule since there are equal numbers of hydroxyl groups on each side of the molecule.

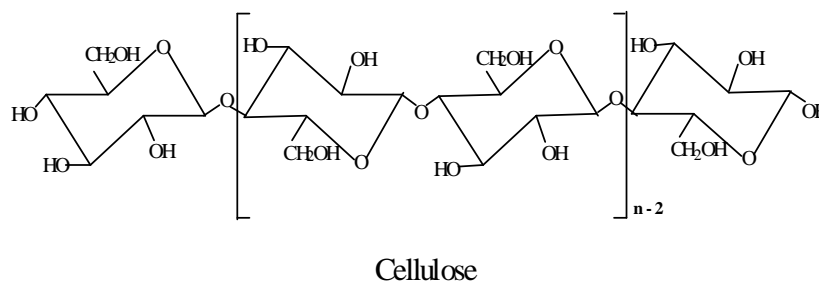


Figure 6: Repeating unit of a cellulose molecule (Zhang & Lynd, 2004).

Researchers have been working on the production of artificial silk from cellulose by forming cellulose derivatives for years. These researches introduced two commonly used routes for the production of fibers: acetate and xanthate esters. Cellulose acetate is a cellulose derivative and soluble in solvents such as acetone and can be spun into fibers. When cellulose is exposed to strong alkali and then treated with  $C_2S$ , xanthate esters of cellulose are formed. Cellulose xanthate is soluble in alkali (aq.) and from this solution filaments and films can be formed. Cellulose xanthate process is the basis of viscose rayon production (Brown, 1982).

Viscose is a regenerated cellulose fiber and it has a high tenacity and high extensibility. It is manufactured from cotton linters or from cellulose obtained from wood pulp. Viscose process requires many steps. Viscose fabrics are less strong than cotton fabrics. Viscose has a very low mechanical strength especially when it is wet.

Viscose is composed of both amorphous and crystalline cellulose with different ratios. The viscose fiber consists of a core surrounded by the mantle, of which crystalline and amorphous cellulose content differs (Figure 7). The crystalline regions in the mantle are smaller and distributed homogeneously throughout the fiber and the core region contains a disordered network formed from bigger crystallites separated by big amorphous regions. The outer region is more ordered and rigid than the core region and accounts for the most of the tensile strength.

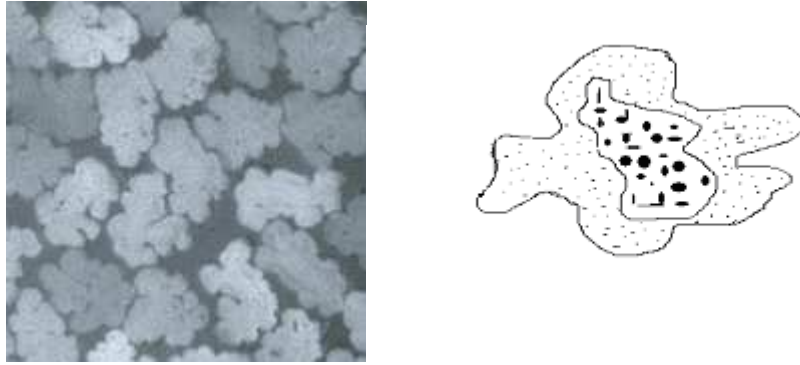


Figure 7: Cross-section of a viscose fiber (Lenzing Viscose® 2,8 dtex) and its schematic representation.

It is known that amorphous cellulose is more prone to attack by cellulases than crystalline cellulose. Crystalline cellulose is more rigid and gives tensile strength along the fiber axis whereas amorphous cellulose is mainly responsible for the flexibility. The loss of tensile strength is most probably due to the loss of highly ordered crystalline structure by the action of cellulases (Lenting & Warmoeskerken, 2001).

#### **1.4. Protein Engineering**

Protein engineering studies have begun in the early 1980s. Chemical modification methods or genetic methods are used to modify the structure of a protein in order to alter its function, activity or stability. This concept uses interdisciplinary approaches, hybrid methods such as use of X-ray crystallography, DNA modification techniques, computer modelling, artificial gene synthesis etc. to achieve protein modification (Bornscheuer & Pohl, 2001; Carter, 1986). Application of protein engineering methods to enzymes has been gaining importance for the last three decades since enzymes are used in many industrial applications. Developments in the area of genetic manipulation allowed researchers to produce those enzymes in homologous or heterologous hosts in larger amounts. Many of the industrial processes still need enzymes with high chemo-regio and stereoselectivity and stability. As a result, screening and production of enzymes with novel properties that can fulfill the needs of these industrial processes is becoming an important issue. This can be achieved through screening of environmental samples and culture collections but high throughput screening methods are needed and this method may not always yield a suitable enzyme. Tailor made biocatalysts can be designed from wild type enzymes via protein engineering using rational design (computer-aided molecular modeling and site-directed mutagenesis) or by directed

evolution of these enzymes (Bornscheuer & Pohl, 2001; Schmidt, Bottcher, & Bornscheuer, 2009; Shao & Arnold, 1996).

#### **1.4.1. Rational Design of Proteins**

Rational design of proteins involves computer aided molecular modelling and site directed mutagenesis of the protein of interest. Computer assisted modelling of biocatalysts have been gaining importance for understanding the underlying physical basis of the structure-function relationships of biological macromolecules. For the last two decades, computer simulation methods have been applied to structural and dynamic studies of many proteins as well as understanding the mechanisms of protein folding and unfolding (Daggett, 2006; De Mori, Meli, Monticelli, & Colombo, 2005; Gu, Wang, Zhu, Shi, & Liu, 2003; Karplus & Sali, 1995; H. L. Liu, Wang, & Hsu, 2003) and recently to the design of enzymes with improved or specific properties (Huang, Gao, & Zhan, 2011) or design of protein inhibitors (Lameira, Alves, Tuñón, Martí, & Moliner, 2011).

Molecular dynamics (MD) is a computer simulation technique that explores protein dynamics in atomic detail (Adcock & McCammon, 2006). Generally all-atom level simulations with longer time scales and higher resolution are preferred in MD simulations since analysis of these give more elaborate information about the energetics and structure of a protein at different temperatures (Beck & Daggett, 2004; Mark & van Gunsteren, 1992). This method, when well designed, can provide detailed information about the protein under study (van Gunsteren & Mark, 1992a, 1992b). We now have a deeper understanding about the concept of unfolded proteins (Floriano, Domont, & Nascimento, 2007). The concept has transformed into a more complex state other than mere unfolding of a polypeptide chain into an extended conformation. The current view states the unfolded state as an ensemble of partially folded conformers of the protein and the denaturing conditions determine the extent of unfolding (Floriano, et al., 2007; Hung, Chen, Liu, Lee, & Chang, 2003). An increase in temperature results in increased intramolecular motions, which in turn causes protein unfolding (H. L. Liu & Wang, 2003a, 2003b). MD simulations of proteins are currently restricted to microseconds (Klepeis, Lindorff-Larsen, Dror, & Shaw, 2009) due to the large computational demands of such simulations. In reality, it has been estimated that half-time required for most proteins for folding is more than 1 millisecond. The rate of unfolding

increases with increased temperatures and it is known that at 225 °C, most proteins unfold in less than 1 nanosecond. MD simulations are generally performed at high temperature because of the large time scales needed for modelling unfolding or folding of proteins (Fersht & Daggett, 2002). It is hypothesized that use of high temperatures in simulations did not change the overall unfolding pathway but just accelerates the kinetics of unfolding (Daggett, 2006; Day, Bennion, Ham, & Daggett, 2002). Additionally, information gathered through application of molecular mechanics simulations can be used to determine the possible sites for directed mutagenesis and the effect of those mutations on the stability of the protein residues. Baysal and Atilgan (2001) introduced a new molecular mechanics approach that involves use of perturbation response theory to study residue stabilities of proteins in a given conformation (Baysal & Atilgan, 2001). In this approach, a perturbation is induced in the form of a displacement on a selected residue of the energy minimized protein followed by energy minimization. The displacement of each residue was recorded in a perturbation response matrix and the stabilities of each residues are calculated from that matrix. They have shown that residue stability had arisen as a tool reflecting the character of the response. By introduction of such perturbations, residues that confer stability and instability to the protein could be identified.

The information about the protein of interest gathered through molecular simulations are used for designing and construction of mutants of the protein with desired properties. Mutations are introduced to DNA encoding these proteins via site directed mutagenesis (Carter, 1986) and computer simulations of these mutants are also performed to evaluate the effect of those mutations on the protein structure. After some random mutagenesis experiments, the mutants with desired properties can be subjected to computer simulations to uncover the the effect of such mutations on the structure and function (Bornscheuer & Pohl, 2001; Shao & Arnold, 1996).

#### **1.4.2. Directed Evolution (Molecular Evolution)**

Directed evolution aka molecular evolution is a very powerful technique which involves either random mutagenesis of a gene of interest or recombination of the gene fragments. the variants produced by either mutagenesis methods are usually screened with high throughput methods to identify and select the desired variant. Directed evolution is developed in 1996 by Francis Arnold (Arnold, 1996) . Arnold used repeated

cycles of mutagenesis and the power of natural selection to force the evolution of the selected protein into a protein with the desired properties. In this method, random mutagenesis is applied and after each cycle of random mutagenesis, selection pressures are applied and only a single variant exhibiting desired improvements is selected and the random mutagenesis procedure is repeated on this variant. This mutagenesis and selection cycles are repeated until obtaining a mutant with desired properties. There are many examples of protein engineering by directed evolution in the literature. You and Arnold (1996) used repetitive rounds of error prone PCR to introduce random mutations and screened the resultant mutant libraries of subtilisin E protein of *Bacillus subtilis*. They obtained a mutant where the total catalytic activity of subtilisin E is significantly enhanced in a non-natural environment, aqueous dimethylformamide (You & Arnold, 1996).

### **1.5. Crosslinked Enzyme Aggregates (CLEA)**

Immobilization of biomolecules on solid supports and crosslinking of enzymes have been gaining importance over the last 40 years. The developments in the biosensor technology and in the enzyme technology enable the researchers to find different supports and crosslinking methods suitable for different applications. Crosslinking of biomolecules such as enzymes offers many advantages. Most important ones are increased stability and durability, multiple use, longer half-life.

No support material is used for crosslinking. Instead, enzymes are crosslinked to each other. Crosslinking of the enzymes causes their aggregation and helps their recovery from the solution. Mostly, two types of crosslinkers are used in this method. Homobifunctional crosslinkers are the ones that bind to the same reactive groups on both sides and heterobifunctional crosslinkers are the ones that have the capability of binding to different reactive groups on each side. Multifunctional crosslinkers are also available for use. Glutaraldehyde is a homobifunctional crosslinker and forms oligoglutaraldehyde in solution (Figure 8).

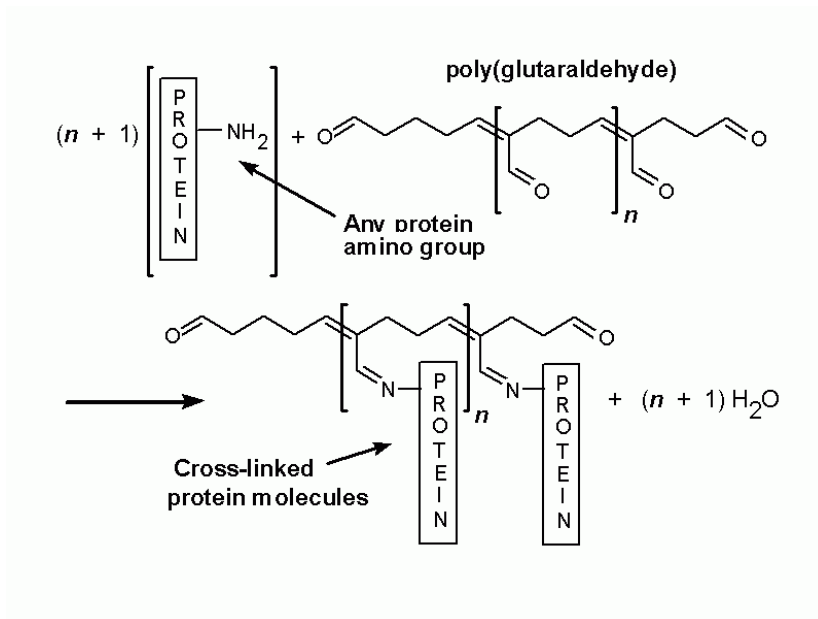


Figure 8: Crosslinking of a protein with a homobifunctional crosslinker, glutaraldehyde (Kiernan, 2000).

It is known that glutaraldehyde reacts with  $\epsilon$ -amino groups on Lysine residues and also N-terminal amino groups (Richards & Knowles, 1968). This reaction is through the double bonds of its oligomeric form. It can not be a single bond since it is very stable and formation of a simple Schiff-base can not provide that stability. Moreover, freshly distilled solutions of glutaraldehyde indicates lower reactivity on proteins. Glutaraldehyde is widely used in crosslinking because of technical ease and versatility of its application.

Many studies propose use of crosslinking agents to enhance stability of the enzyme but they seem to have adverse impact on the activity since they target certain types of aminoacids that may also exist in the vicinity of the active site (Busto, Ortega, & Perez-Mateos, 1997; Yuan, Shen, Sheng, & Wei, 1999). Crosslinking with agents like glutaraldehyde is known to reduce enzyme activity due to unspecific binding to Lysine residues (Busto, et al., 1997).

Crosslinked enzymes form a large, three-dimensional complex structure. Since the crosslinking attaches all the enzymes together, reduced activity or stability due to the steric hindrance is expected. Introduction of spacer molecules or proteins such as Bovine Serum Albumin (BSA) may be a solution to the close proximity problems or



creation of hotspots for directed crosslinking away from the active site may be another way to address this problem.

Nowadays, the activity losses as a result of crosslinking is overcome by the introduction of Crosslinked enzyme crystals (CLEC) and Crosslinked enzyme aggregates (CLEA) technologies. Since it is hard to get enzymes as crystals CLEC is not used very much but CLEA technology is much more simple and proven to cause hyperactivation of glucose oxidase, laccase, lipase enzymes (López-Serrano, Cao, van Rantwijk, & Sheldon, 2002).

CLEAs are prepared by precipitating the enzyme molecules by a polar precipitant solution such as ammonium sulphate, ethyl lactate, PEG, acetone, tert-butyl alcohol etc. and then crosslinking these aggregates with a suitable crosslinker (mostly glutaraldehyde). It is shown that formation rate of aggregates increases as the polarity of the solvent increases. Use of CLEAs is a universal and cheap alternative to other crosslinking methods because pure proteins with enhanced activity and having higher protein ratios per volume is reached.

## **CHAPTER 2**

### **2. PURPOSE**

Viscose fabrics are more prone to pilling than any of the other fabric types. There are no cellulase formulations that effectively remove the microfibrils that cause pill formation on the surface of the viscose fabrics. Aggressive action of the cellulases causes loss of fabric strength due to the damage in the highly ordered crystalline regions of the viscose fibers. One solution to this problem is the design and production of enzymes with increased molecular weights so that aggressive action of the cellulases would be limited to the fabric surface. The aim of this study is the design and production of cellulases and cellulase formulations that can alleviate the problem of pilling and loss of tensile strength in viscose fabrics and evaluation of their effects on viscose fabric properties. For this purpose, throughout this study, both protein engineering and chemical modification methods were used separately and in combination to obtain enzymes with increased molecular weights.

## CHAPTER 3

### 3. MATERIALS AND METHODS

#### *3.1. Molecular Modeling*

##### *3.1.1. Molecular Models*

A novel molecular mechanics approach introduced by Baysal and Atilgan was used to determine the possible loop insertion regions. This approach involves the use of perturbation response theory. Stabilizer motifs on EGI were identified using MM simulations with the application of perturbation response theory and flexibility analysis of EGI with FIRST (Floppy Inclusions and Rigid Substructure Topography) software now located in Flexweb (<http://flexweb.asu.edu/software/first/>). Two potential loop insertion sites were determined based on these data. These sites were between 112<sup>th</sup> and 113<sup>th</sup> residues and between 155<sup>th</sup> and 156<sup>th</sup> residues.

All loop models were constructed using Modeller 8v2 (Sali & Blundell, 1993) using endoglucanase 1 native 3-D structure (PDB code: 1EG1) as a template. Ten aminoacid long loops composed of four different compositions of Lysine and Glycine residues were introduced into two flexible regions of EGI. Glycines were introduced for spacing and flexibility, Lysines were introduced as the crosslinking spots. L1, L2, L3, L4 were introduced between 155<sup>th</sup> and 156<sup>th</sup> residues and L5, L6, L7 were introduced

between 112<sup>th</sup> and 113<sup>th</sup> residues of the protein. L1 and L5, L2 and L6, L3 and L7 have the same loop sequence and composition.

### ***3.1.2. Molecular Dynamics (MD) Simulations***

MD simulations were used to analyze the effects of the proposed mutations (L1-L7) on the native enzyme structure (EGI). NAMD/VMD software package was used for MD simulations (Phillips, et al., 2005). Simulations were performed inside 6 Å<sup>3</sup> waterboxes under periodic boundary conditions and they were run at different temperatures: 300 °K, 400 °K, 450 °K, 500 °K in order to see the effect of temperature on the structure of the enzymes. 2 fs timestep was used and data collection was done every 2 ps. Structures were minimized 10000 steps using conjugate gradient (CG) method before the simulations.

Simulations were performed at higher temperatures to speed up the unfolding kinetics. Unfolding behavior of the enzyme was reflected by the change in root mean squared deviation (RMSD) and radius of gyration (RGYR) values with the increase in temperature. These values were calculated from the trajectory files using VMD and Carma (Glykos, 2006) programs, respectively. Both RMSD and RGYR are indicators of the degree of dissimilarity for the structures at the end of the simulation at the simulation temperature with respect to their initial structures. RGYR indicates how much the structure spreads out from its center. Size and shape were both taken into account in RGYR calculations. Shorter simulations (4 ns) were performed for EGI and its all loop mutants. Best loop mutant (EGI\_L5) was selected for further detailed analysis. In order to confirm the unfolding behavior of the native and mutant enzyme properly, longer simulations (10 ns) were carried out at 300 °K and 450 °K. Effect of the proposed mutation on the active site was analyzed by calculating active site residue distances from trajectory files with VMD programme.

### ***3.1.3. Molecular Mechanics (MM) Simulations***

Perturbation response analysis by MM simulations method developed by Baysal and Atilgan was also used to study the effect of loop insertion on stabilizer motifs of the enzyme. The native enzyme EGI is compared to its loop inserted mutant EGI\_L5 utilizing MM procedure. 10000 step initial minimization was applied to both enzymes with CG method until the gradient tolerance was 0.001. C $\alpha$  of each residue was pulled

along a distance. New coordinates were fixed. The energy of the system was minimized for 400 steps more with CG method. This was repeated for each residue in the sequential order for the model protein. The pulling was used as the perturbation and the displacement of each residue in response to such perturbation is recorded in a perturbation-response matrix  $L$ .

Protein stability is defined by  $\lambda_i = \sum_j L_{ij} / \sum_j L_{ji}$  where  $\lambda_i$  is the ratio of the total amount of change that may be induced on the same residue by equal amounts of perturbations inserted in the rest of the protein, to the overall ability of a given residue to induce change in the rest of the protein. The free energy associated with the residue stability is defined by  $\Delta G_\lambda = -RT \ln \lambda$  (Baysal & Atilgan, 2001). From this perspective, idle loops will have  $\Delta G_\lambda < 0$ , the residues that contribute to stability and function will have  $\Delta G_\lambda > 0$ , functional loops will have  $\Delta G_\lambda \approx 0$ , hydrophobic core residues that have insignificant contributions to stability or function will also have  $\Delta G_\lambda \approx 0$ .  $\Delta G_\lambda / RT$  values of each residue were calculated from perturbation-response matrices of EGI and EGI\_L5. The effect of the proposed loop insertion on the stability of the active site of the enzyme was analyzed.

### **3.2. Microorganisms, enzymes and chemicals**

*Escherichia coli* Top10 F was used as host for the propagation of the vectors and subcloning. *P. pastoris* KM71H (aox1::ARG4, arg4) (Invitrogen, San Diego, USA) was used for recombinant protein expression. pPICZ $\alpha$ A vector (Invitrogen, San Diego, USA) was used for cloning and protein expression. Pfu Polymerase and Rapid Ligation Kit (Fermentas) were used for cloning purposes. Taq Polymerase (Qiagen) was used for colony PCR. All of the contained chemicals were of analytical grade.

### **3.3. Site directed mutagenesis**

Overlap PCR extension method was applied to introduce the loop mutation. Overlap PCR extension primers for preparation of *pPiczaA\_eglI\_L5* (Figure 9, Table 2) and *pPiczaA\_eglI\_BC* (Figure 10, Table 3) were designed according to literature (Vallejo, Pogulis, & Pease, 1994). L5 loop mutation was introduced between 112<sup>th</sup> and 113<sup>th</sup> aminoacids. The fidelity of the constructs was confirmed by sequencing.

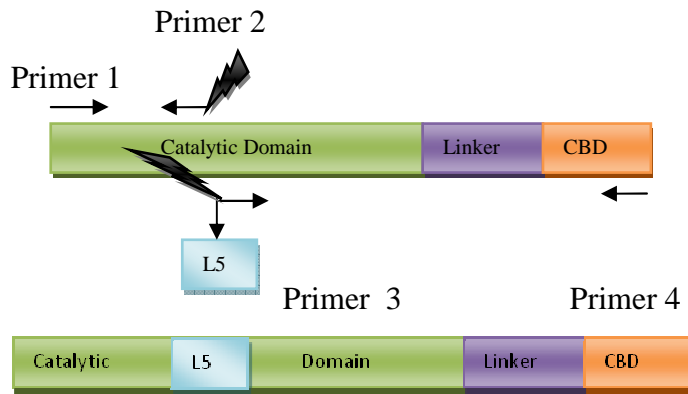


Figure 9: Overlap extension PCR schematics for *egl1\_L5*. *egl1\_L5* mutant gene (bottom drawing) prepared from *egl1s* ( top drawing). Overlap extension primers were denoted with numbers. Added loop domain, L5 was shown as flashes at the end of each overlap extension primer.

Table 2: Overlap PCR extension primers designed for *egl1\_L5* gene.

.	Primer Sequence	Notes
Primer 1	5'CCG <u>GAATTCC</u> CAGCAACCGGGTACCAGCACC 3'	<u>EcoRI</u> restriction site
Primer 2	5' <u>ACCTTCTTACCCTTCTTCTTACCCTTCTT</u> GTCCA AGTACAATCTTG 3'	<u>L5</u> sequence
Primer 3	5' <u>AAGAAGGGTAAGAAGAAGGGTAAGAAAGGT</u> TCC GACGGTGAATACGGT 3'	<u>L5</u> sequence
Primer 4	5'CCG <u>TCTAGAG</u> CAAGGCATTGCGAGTAGTAG 3'	<u>XbaI</u> restriction site

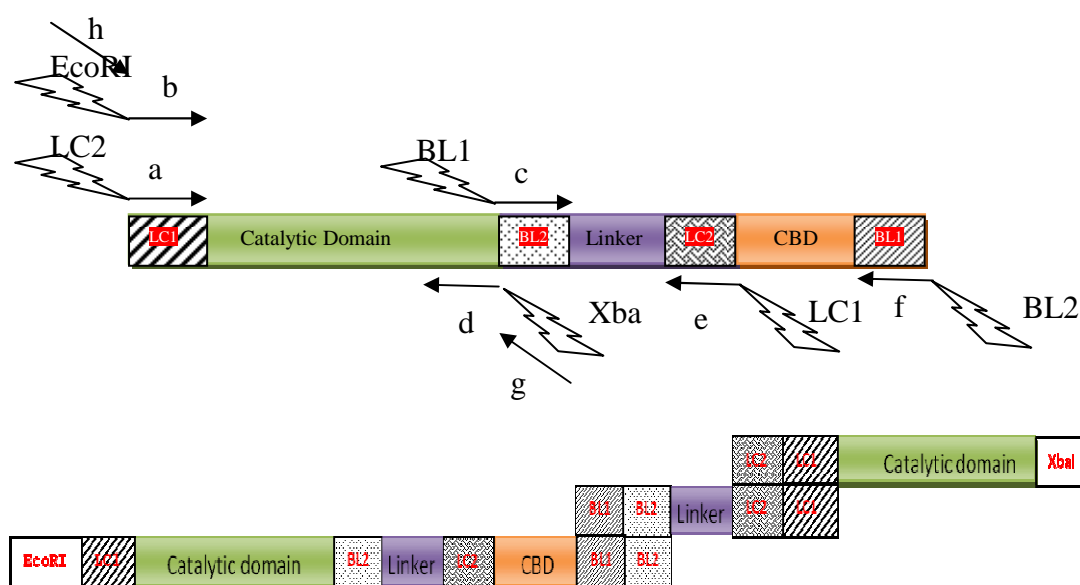


Figure 10: Overlap extension PCR schematics for *egl1\_BC*. *egl1\_BC* mutant gene (bottom drawing) prepared from *egl1s* (top drawing). Overlap extension primers, overlapping regions were denoted with letters. Added overlapping regions were shown as flashes at the end of each overlap extension primer.

Table 3: Overlap PCR extension primers designed for *egl1\_BC* gene.

	Primer Sequence	Notes
<b>Primer a</b>	5' <u>TCCATCCTGTCAGCAACCGGGTACCT</u> 3'	LC2 region
<b>Primer b</b>	5' GGGCCATCCAGAATT <u>CAGCAACCGGGTAC</u> CT 3'	<i>EcoRI</i> restriction site
<b>Primer c</b>	5' <u>GCAATGCCTT CCACCACCACCTG</u> 3'	BL1 region
<b>Primer d</b>	5' CGTCATCGGCT <u>CTAGAAGCAGTAGAGTTTGT</u> A 3'	<i>XbaI</i> restriction site
<b>Primer e</b>	5' <u>CCGTTGCTGACAGGATGGAGAGGAA</u> 3'	LC1 region
<b>Primer f</b>	5' GTGGTGGTGGAAAGGCATTGCGAGTAG 3'	BL2 region
<b>Primer g</b>	5' CGTCATCGGCT <u>CTAGA</u> 3'	<i>XbaI</i> restriction site
<b>Primer h</b>	5' GGGCCATCCAGA <u>ATTTC</u> 3'	<i>EcoRI</i> restriction site

### 3.4. Cloning

A codon optimized (for *Pichia pastoris*) endoglucanase 1 synthetic gene (*egl1s*) carrying *EcoRI* and *XbaI* restriction sites on each arm was obtained from GeneART. Protein sequence of the synthetic gene is identical to *Trichoderma reesei* endoglucanase 1 (GenBank accession no. **M15665**). *EcoRI* and *XbaI* sites were added to the gene flanking regions. For cloning *egl1*, *egl3*, *cbh1* and *cbh2* genes, total RNA from *Trichoderma reesei* QM9414 strain was extracted using RNAeasy RNA isolation kit (Qiagen) according to the manufacturer's instructions. cDNA was synthesized from

total RNA using RT-PCR. *EcoRI* and *XbaI* sites were added to the gene flanking regions of *egl1*, *egl3*, *PstI* and *NotI* sites were added to the gene flanking regions of *cbh1*, *EcoRI* and *NotI* sites were added to the gene flanking regions of *cbh2* during this RT-PCR reaction. The synthetic gene endoglucanase 1 (*eglls*) and the native cDNA's were double-digested with suitable restriction enzyme couples and ligated into suitable multiple cloning site of pPiczaA (*eglls*, *egl1*, *egl3*, *cbh2*) or pPiczaB (*cbh1*) vectors using Rapid Ligation Kit (Fermentas). All genes were subcloned in *E.coli* Top10 F or XL1-Blue cells. *E.coli* cells were cultured on low salt LB plates in the presence of 25 µg/ml zeocin. Zeocin positive colonies were selected and colony PCR with 5' AOX (5'-GACTGGTTCCAATTGACAAGC-3') and 3' AOX (5'-GCAAATGGCATTCTGACATCC-3') primers was performed. Colony PCR positive colonies were selected and recombinant plasmids were isolated using MiniPrep Kit (Qiagen). Fidelity of the constructs were confirmed by sequencing.

### **3.5. Transformation and Screening**

The recombinant plasmids and empty vector pPiczaA were linearized with either *SacI* or *PmeI* or *BstXI* before transformation according to their restriction sites. Prepared plasmids and their linearization sites were shown in Table 4. This process resulted in the stable integration of one or multiple copies of the linearized vector at 5' AOX1 chromosomal locus of *P. pastoris* KM71H by homologous recombination. All the obtained transformants were Mut<sup>s</sup> (slow methanol utilizing phenotype). Competent *P. pastoris* KM71H cells were prepared according to a procedure combining chemical transformation and electroporation (Wu & Letchworth, 2004). ~1µg linearized recombinant plasmid was mixed with competent KM71H cells. The mixture was immediately transferred to a pre-chilled 0.2cm electroporation cuvette and incubated on ice for 5 minutes. About 1 ml of ice-cold 1M sorbitol was immediately added to the cuvette after electroporation. The charging voltage, capacitance, and resistance were 1.5 kV, 25F, and 200 Ω, respectively. The transformation mixture was spread onto YPD plates containing 100 µg/ml zeocin. The plates were incubated at 30 °C until the appearance and growth of colonies (about three days). Zeocin positive colonies were selected and colony PCR with 5' AOX and 3' AOX primers was performed. Colony PCR positive colonies were selected.



Multicopy transformants were selected on BMM-agar plates containing blue colored Azo- carboxymethylcellulose (Azo- CMC) as a substrate. More active transformants expressing the enzyme were selected according to the relative radii of the clear zones around the colonies. When the enzyme was active, it degraded Azo-CMC and clear zones were produced around the colonies as a result of enzymatic hydrolysis. Selected clones and their notations used in the research were shown in Table 4.

Table 4: Plasmid names and linearization sites for all genes.

Plasmid	Restriction Enzymes for Linearization	Clones
<b>pPiczaA</b>	<i>PmeI</i>	A4
<b>pPiczaA-egl1</b>	<i>SacI</i>	E12, X11
<b>pPiczaA-egl1n</b>	<i>PmeI</i>	C5
<b>pPiczaA-egl1_L5</b>	<i>SacI</i>	D5
<b>pPiczaA-egl1_BC</b>	<i>SacI</i>	F7, F8
<b>pPiczaA-egl3</b>	<i>SacI</i>	C13, Z10
<b>pPiczaB-cbh1</b>	<i>BstXI</i>	X3

### 3.6. Copy Number Determination

Copy number of *egl1s* and *egl1* expression cassettes integrated into the *Pichia pastoris* was determined by quantitative real time PCR. SYBR Green PCR Master Mix (Fermentas) was used in a BioRAD iCycler according to the manufacturer's instructions. Genomic DNA was extracted from stable transformants and 100 ng of genomic DNA was used as a template in the quantitative real time PCR reaction. Primers for the promoter region of AOX gene were used for quantitative real time PCR. Primers for *Pichia pastoris* ARG4 gene were used for normalization. It has been known that *Pichia pastoris* harbors only a single copy of ARG4 gene. Following cycling parameters and primers were used for the amplification: Forward ARG4 primer, 5' TCCTCCGGTGGCAGTTCTT 3'; reverse ARG4 primer, 5' TCCATTGACTCCCGTTTTGAG 3'; forward AOX promoter primer, 5'ACATCCACAGGTCCATTC 3'; reverse AOX promoter primer, 5'GGTGTTAGTAGCCTAATAGAAG 3'. Initial denaturation at 95 °C for 10 min.; 40 cycles of denaturation at 95 °C for 15 sec., annealing at 60 °C for 30 sec., extension at 72°C for 30 sec.

### 3.7. Small Scale Expression of recombinant *Pichia pastoris* strains

The colonies were inoculated into 50 mL BMG medium (100mM potassium phosphate, pH 6.0, 1.34% YNB,  $4 \times 10^{-5}$  % biotin and 1% glycerol) and grown (250rpm) at 30 °C overnight in 250 ml baffled shake flasks. When OD<sub>600</sub> reached 10 units/ml, the cells were collected by centrifugation (3000×g, 5min). The cell pellet was resuspended in BMM medium (BMG with 0.5% (w/v) methanol instead of 1% glycerol) with a starting OD<sub>600</sub> of 30 units/ml. The culture was grown for ~ 120 hours at 30 °C. Methanol was added to a final concentration of 10 g/L at every 24 hours. Cell culture supernatants were collected every 24 hours.

### ***3.8. Bioreactor Cultivations of recombinant Pichia pastoris strains***

*Pichia pastoris* clones harboring EGI, EGI<sub>n</sub>, EGI\_L5, EGI\_BC, EGIII, CBHI, CBHII were grown in a 7.5L fermenter (BioFlo 110, New Brunswick Scientific) with a starting volume of 2L at 28 °C and pH 5 with feed rates of 18 ml/h/L glycerol and 1-12 ml/h/L methanol. EGI\_BC subjected to three different fed-batch fermentations at pH 5, 29 °C; at pH 7, 29 °C and at pH 5, 25 °C. Three different conditions were applied to prevent degradation of the product. Invitrogen Corp.'s fermentation medium recipe was used for fed-batch fermentations along with PTM1 trace metal solution. Glycerol was used as the sole carbon source throughout glycerol batch and glycerol fed-batch phases of the fermentations. Methanol was used as an inducer of the AOX promoter during methanol fed-batch phase of the fermentations. Methanol levels were monitored using a specific methanol probe (Raven Biotech). Fermentation products were filtered, buffer exchanged and concentrated using Sartocon Micro and Ultrafiltration System (Sartorius-Stedim). Sartocon Slice 200 HydroSart membranes with 0.45 micron cutoff and Polyether sulfone (PES) membranes with 100 kDa and 10 kDa cutoffs were used. Enzymes were either lyophilized after filtration or frozen at -20 °C for long-term storage.

### ***3.9. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Zymogram Analysis***

Collected cell culture supernatants were run on an SDS-PAGE gel with 5% stacking gel and 12% separating gel. About 15–20 µl of the supernatant was loaded into each well of the gel. After electrophoresis, the gel was stained with Coomassie Brilliant Blue R-250. For the zymogram analysis native SDS-PAGE was performed. The gels were washed with 2.5 % (v/v) TritonX-100- 50 mM sodium acetate solution at pH 4.8 three

times for 10 minutes. Then, the gels were washed with 50 mM sodium acetate buffer at pH 4.8 three times for 10 minutes. Zymogram analysis was performed using 4-MUC (4-Methylumbelliferyl beta-D-cellobioside) as a substrate. Activity of the enzymes were visualized and documented under UV-light. After zymogram analysis, the gels were stained with Coomassie Brilliant Blue R-250. All gel photographs were documented using Gel-Doc (BioRad).

### ***3.10. Purification of Recombinant Proteins***

EGI, EGI<sub>n</sub>, EGI<sub>BC</sub> and EGI<sub>L5</sub> were purified using regenerated amorphous cellulose (RAC) as an affinity chromatography matrix. Batch affinity purifications of both enzymes were obtained. Hong et al.'s purification methodology was employed with slight modifications (Hong, Ye, Wang, & Zhang, 2008). Avicel Ph105 was used for the synthesis of regenerated amorphous cellulose. Purification was performed at room temperature. Recombinant cellulases were eluted with ethylene glycol and ethylene glycol was further removed and exchanged with 50 mM Sodium acetate buffer (pH 5) due to its interference with BCA protein assays by ultrafiltration through 10 kDa cutoff Vivaspin 500 membrane spin filters (Sartorius-Stedim).

### ***3.11. Protein Assays***

Protein concentrations were determined using BCA Protein Assay Reagent (Pierce) according to manufacturer's instructions. Bovine serum albumin was used as the protein standard.

### ***3.12. Enzyme Assays***

All CMC activity assays were performed in triplicates with a standard deviation of below 10%. All 4-MUC activity assays were performed in duplicates or triplicates with a standard deviation below 10%.

#### ***3.12.1. Effect of Temperature on Enzyme Activity***

Activity of each enzyme at different temperatures (15°C -95 °C) were determined by 3,5-Dinitrosalicylic acid (DNS) method against 0.5 % CMC (w/v) in 50 mM sodium acetate buffer (pH 4.8) (Ghose, 1987). Each enzyme and substrate was

preincubated for 5 minutes at each assay temperature separately. Enzyme and the substrate were incubated at the assay temperature for 10 minutes. Reducing sugars produced were measured at 550 nm. Glucose was used as a standard. Residual enzyme activity was calculated by taking the maximum activity of the enzyme at the determined temperature as 100 %.

### ***3.12.2. Effect of pH on Enzyme Activity***

Activity of each enzyme at different pHs (pH 3-7) were determined using 3,5-Dinitrosalicylic acid (DNS) method against 0.5 % CMC (w/v) in different pH buffers for 10 minutes at 55 °C. Reducing sugars produced were measured at 550 nm. Glucose was used as a standard. Residual enzyme activity was calculated by taking the maximum activity of the enzyme at the determined pH as 100 %.

### ***3.12.3. 4-MUC Assays***

4-MUC assay was performed for fermentation products according to (Chernoglazov, Jafarova, & Klyosov, 1989). Fermentation samples were incubated with 0.5 mg/ml 4-MUC in 50 mM sodium acetate buffer at pH 4.8. Kinetic analysis was performed at 25 °C, 30 °C or 45 °C for 30 minutes. 4-Methylumbelliferone (4-MU) was used as the standard. Liberated 4-MU was measured with a fluorescence spectrophotometer with excitation at 363 nm and emission at 435 nm.

### ***3.12.4. Stability assays***

Enzyme stability was determined by first incubating each enzyme at 100 °C for 10 minutes. Enzymes were then chilled on ice for 10 minutes. Standard CMC assay was performed to 100 °C incubated and unincubated enzyme samples at 55 °C and pH 4.8 for 10 minutes. The stability of each enzyme was calculated with respect to its normal activity in the form of retained activity.

Residual enzyme activity at 50 °C was determined by incubating both of the enzymes at 50 °C for 0h to 72h and assaying the enzyme activities against 1% CMC

(w/v) at 50 °C and pH 4.8 for 10 minutes. Activity of the enzymes with no incubation (time 0) at 50 °C were taken as 100 %. At least two independent activity tests were carried out for each incubation. At each incubation time, recombinant enzymes were assayed in triplicates with a standard deviation below 10 %.

Residual enzyme activity at 70 °C was also determined by incubating both of the enzymes at 70 °C for 0h to 2h and assaying the enzyme activities against 1% CMC (w/v) at 50 °C and pH 4.8 for 10 minutes. Activity of the enzymes with no incubation (time 0) and assayed at 50 °C were taken as 100 %. Enzyme activity was determined in triplicates with a standard deviation below 10 %.

### ***3.13. CLEA Preparation***

#### ***3.13.1. CLEA preparation from commercial cellulase***

CLEA of Gempil 4L was prepared according to Schoevaart et al. with slight modifications (Schoevaart, et al., 2004; Sheldon, Schoevaart, & Langen, 2006). Acetone was used as the precipitant solution. 200 ml and 500 ml of the enzyme solution was added to an Erlenmayer flask with a magnetic stirrer bar (batch 2 and batch 3 CLEA). CLEA of Gempil 4L was prepared at room temperature with constant mixing. 1800 ml and 4500 ml of the precipitant solution containing glutaraldehyde as the crosslinker at pH 7.3 was added drop by drop to the enzyme mixture, respectively. The suspensions were stirred for 1 hour. The reaction was quenched by the addition of 200 ml and 500 ml of 1M Tris solution at pH 8, respectively. The suspensions were vacuum filtered through 10 µm metal filter and washed with excess amounts of 0.1 M Potassium phosphate buffer at pH 7.3. As a final step, CLEA particles of Gempil 4L were washed with acetone and dried overnight at room temperature. Dried CLEA particles were ground using TissueLyzer (Qiagen) apparatus for 1 minute at a frequency of 1/30 (1/sec). The ground CLEA particles were separated based on their sizes by further filtering them through metal filters of different sizes ( 10 µm, 25 µm, 33 µm, 45 µm, 77 µm, 154 µm, 288 µm, 1980 µm).

#### ***3.13.2. CLEA preparation from EGI and EGI\_L5***

CLEA of EGI and EGI\_L5 weren also prepared according to Schoevaart et al. with slight modifications (Schoevaart, et al., 2004; Sheldon, et al., 2006). Acetone was used

as the precipitant solution. 20 ml and 100 ml of the EGL\_L5 solution were added to Erlenmayer flasks with magnetic stirrer bars. CLEA of Gempil 4L was prepared at room temperature with constant mixing. 180 ml and 900 ml of the precipitant solution containing glutaraldehyde as the crosslinker at pH 7.3 was added drop by drop to the enzyme mixtures, respectively. The suspensions were stirred for 1 hour. The reactions were quenched by the addition of 20 ml and 100 ml of 1M Tris solution at pH 8. The suspensions were vacuum filtered through 10 µm metal filters and washed with excess amounts of 0.1 M Potassium phosphate buffer at pH 7.3. As a final step, CLEA of the recombinant enzymes were washed with acetone and dried overnight at room temperature. Dried CLEA particles were ground using metal steel balls and vortexing.

### ***3.14. Enzymatic Biofinishing of Viscose Fabrics***

100 % viscose supreme (single Jersey) knitted fabric kindly provided by Denge Chemicals used for biopolishing. It had a fabric density of 137 g/m<sup>2</sup>. The fabric was bleached and primary fibrillation was performed. 100% viscose knitted fabric was used for biopolishing with CLEAs.

Biopolishing of viscose fabric with enzyme samples was performed in shaker incubators under constant shaking (250 rpm) at 55 °C, pH 4.8 in 0.05 M NaOAc buffer for 2 hours in 1 L Erlenmayer flasks. Liquor ratio of 1:9 was used in all tests. A special plastic container with a embroidery hoop like apparatus was prepared for CLEA biopolishing experiments. This apparatus acted as a holder for the fabric. A liquor ratio of 1:50 was used with the apparatus and the container housing the apparatus was put in a shaker incubator for biopolishing under constant shaking (150 rpm) at 55 °C, pH 4.8 for 2 hours. A hybridization chamber and its bottles were also used for biopolishing experiments under constant rotation (15 rpm) at 55 °C, pH 4.8 for 2 hours.

180 µl of the commercial enzyme (Gempil 4L) was used for biopolishing of 1 g viscose fabric in all experiments. Quantity of the modified enzymes and recombinant enzymes used for biopolishing were determined by using an activity equivalent of the modified or recombinant enzyme with respect to Gempil 4L. That is the fraction of the modified enzyme exhibiting the same activity as the native enzyme according to the CMC assay was used.

### ***3.15. Light Microscope and SEM Characterization***

Enzyme treated and untreated fabric samples were analyzed with a digital light microscope and their photographs were taken under two different magnifications (~15X and ~200X).

SEM characterization of the enzyme treated, modified enzyme treated and untreated fabric and fibers was performed with Gemini Supra 35VP without coating and under low accelerating voltage (0.9 kV) and with a working distance of 4 mm using the Inlens detector .

### ***3.16. Pilling Test***

All the tests were performed in Denge Chemicals Physics Laboratory. Martindale 2000 pilling machine was used at 125 to 2000 rpm. The reference photographs used were EMPA Standart SN 198525 K3. The photographs were evaluated according to AATCC (Association for American Textile Chemists and Colorists) standards with eye examination. All the values are the weighted averages of the five measurements. For pilling measurements, a five-point evaluation system is used. 1 indicates intense pilling and 5 indicates no pilling.

### ***3.17. Bursting Strength Test***

Bursting strength tests for enzyme treated and untreated viscose fabrics were performed in James H. Heal testing machine according to AATCC. Quadruple measurements were taken for each sample fabric and the end values were the weighted averages of those.

## CHAPTER 4

### 4. RESULTS

#### *4.1. Modelling and production of Trichoderma reesei endoglucanase 1 and its mutant in Pichia pastoris*

##### *4.1.1. Molecular Modelling*

Molecular dynamics simulations of all loop models constructed with Modeller 8v2 were analyzed and RMSD (Root Mean Squared Deviation) and RGYR (Radius of Gyration) of the simulated structures were calculated from trajectory files as a function of simulation time. RMSD and RGYR of all loop mutants and EGI derived from 4 ns long MD simulations at 450 °K were shown in Figure 11 and 12, respectively.

According to RMSD analysis, EGI\_L5 loop mutant was chosen as a candidate for MM and longer MD simulation studies because of its trajectory's similarity to EGI structure during 4 ns of MD simulations. EGI\_L5 exhibited slightly better properties in terms of RMSD in comparison to all other mutants in the last 3000 ps of MD simulations. At high temperature while rest of the structures' RMSD from the original structure were increasing, EGI\_L5's RMSD was lower than all other loop mutants at 450 °K, meaning the loop insertion L5, rendered the enzyme more stable with respect to other loop mutations. RGYR of all loop mutants and EGI did not follow a similar pattern as the RMSD data. RGYR calculations indicated that on the overall, EGI and EGI\_L2 structures were more compact than all other structures. But EGI\_L2 structure performed poorly in terms of RMSD from its original structure and also with respect to



EGI. 10 ns long MD simulations of EGI and EGI\_L5 at 300 °K (Figure 13) indicated an increase in the RMSD of EGI\_L5 with respect to EGI as expected and at 450 °K (Figure 11, small figure inside) the RMSD of C $\alpha$  of EGI\_L5 have exhibited a slight increase in comparison to EGI structure, thus a slightly more loose structure. At 300 °K both EGI and EGI\_L5 structures have shown no signs of unfolding. Moreover, active site residue distances calculated from 4ns long MD simulations have exhibited that the distance between active site Glu205 (196 in original structure) and His221 (212 in original structure) residues have increased by 0.554 Å and the distance between active site Asp207 (198) and Glu210 (201) residues have increased by 0.21 Å on the average in EGI\_L5 with respect to EGI. The distance between Asp207 (198) and His221 (212) was not affected at all. Active site remained almost intact during 4 ns MD simulation for both structures (Figure 14). Superimposed EGI and EGI\_L5 structures (energy minimized) were shown in Figure 15.

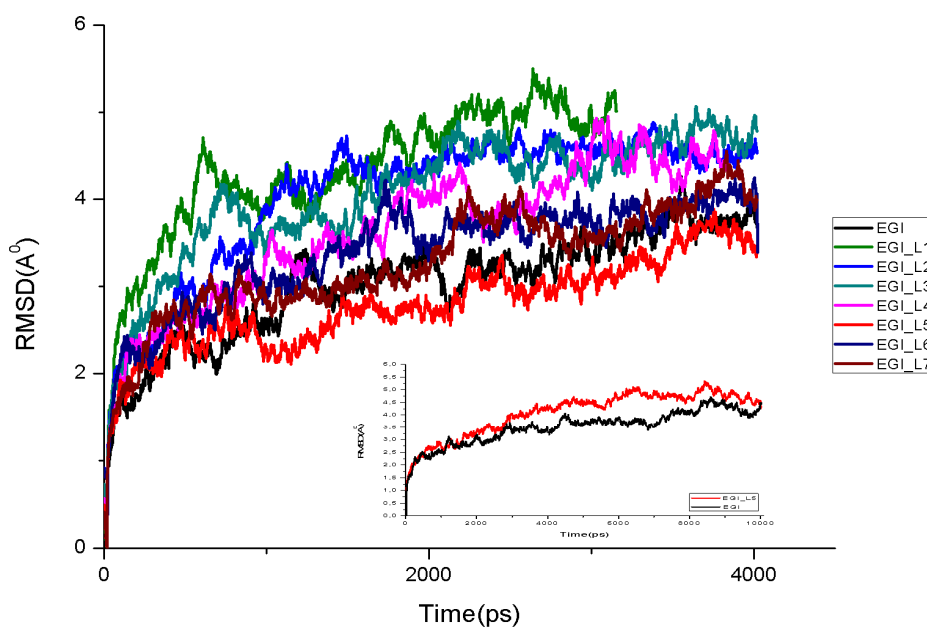


Figure 11: RMSD (Å) of EGI, EGI\_L1, EGI\_L2, EGI\_L3, EGI\_L4, EGI\_L5, EGI\_L6, EGI\_L7 along simulation time (ps) during 4ns MD simulations at 450 °K.

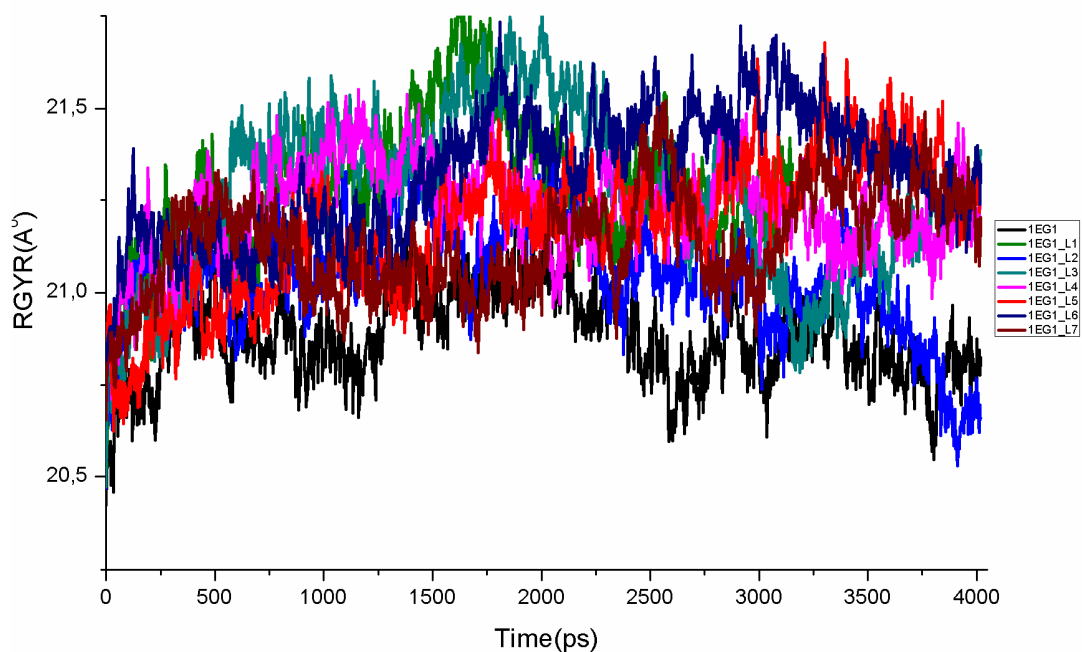


Figure 12: RGYR (A<sup>2</sup>) of EGI, EGI\_L1, EGI\_L2, EGI\_L3, EGI\_L4, EGI\_L5, EGI\_L6, EGI\_L7 along simulation time (ps) during 4ns MD simulations at 450 °K.

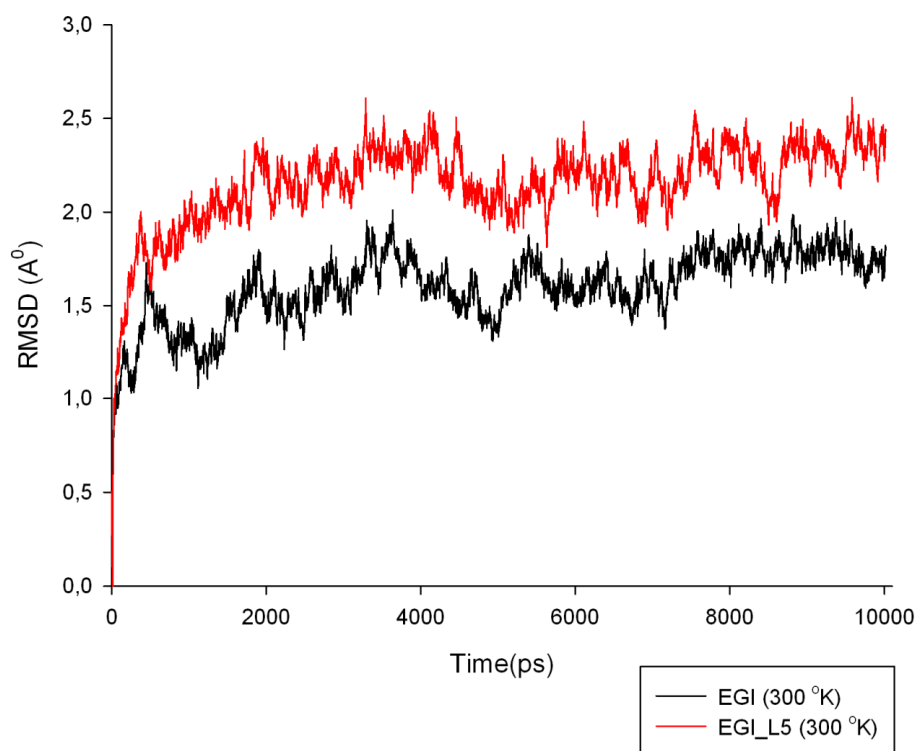


Figure 13: RMSD (A<sup>0</sup>) of EGI vs EGI\_L5 along simulation time (ps) during 10 ns MD simulations at 300 °K.

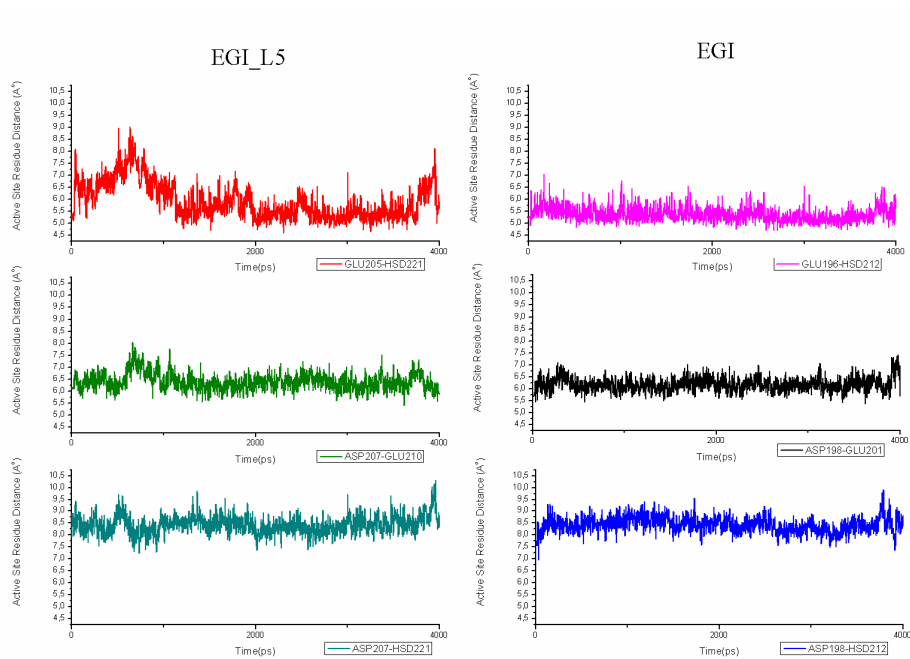


Figure 14: Distance of active site residues to each other during 4 ns MD simulations of EGI and EGI\_L5.

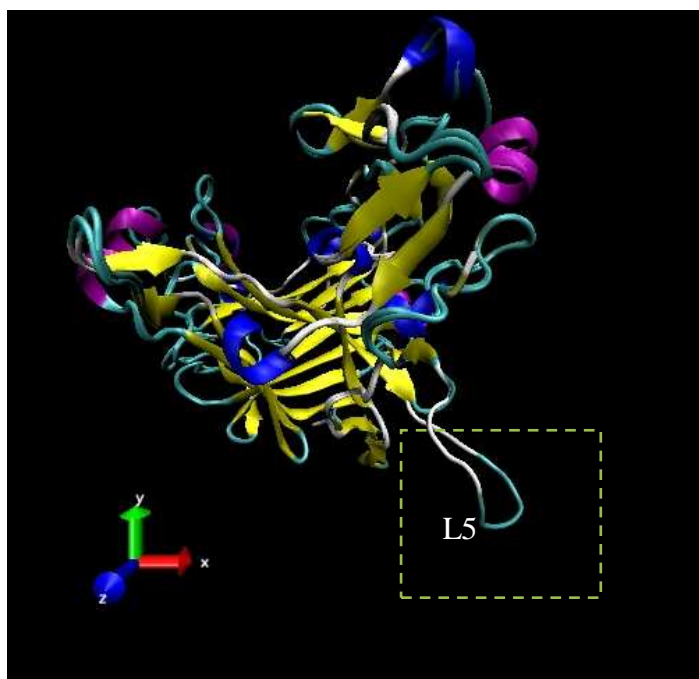


Figure 15: Superimposed structures of EGI and loop mutant EGI\_L5. L5 was inserted between residues 112<sup>th</sup> and 113<sup>th</sup> of EGI.

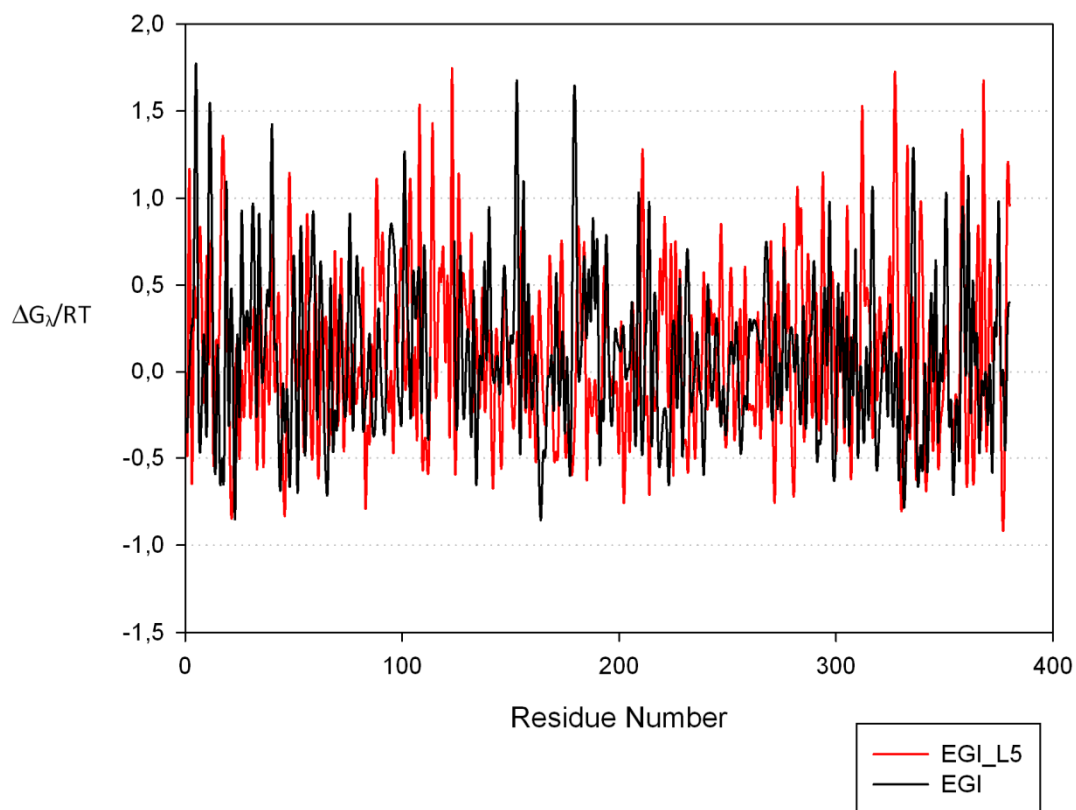


Figure 16: EGI and EGI\_L5 stability coefficients calculated for each residue from molecular mechanics simulations. Stability coefficient  $> 0$  indicates a less rigid structure whereas stability coefficient  $< 0$  indicates a more rigid structure.

Molecular Mechanics simulation results of EGI and EGI\_L5 are shown in Figure 16 and 17. Stability coefficient  $>0$  indicated a less rigid, more flexible structure. Introduction of a ten aminoacid loop composed of Lysine and Glycine between 112<sup>th</sup> and 113<sup>th</sup> residues kept the active site more intact since the stability coefficient was more negative in loop inserted structure compared to the native EGI. The flexibility around the active site decreased in the loop mutant, probably giving the enzyme a slightly more stable structure. The flexibility of the loop region was higher as expected.

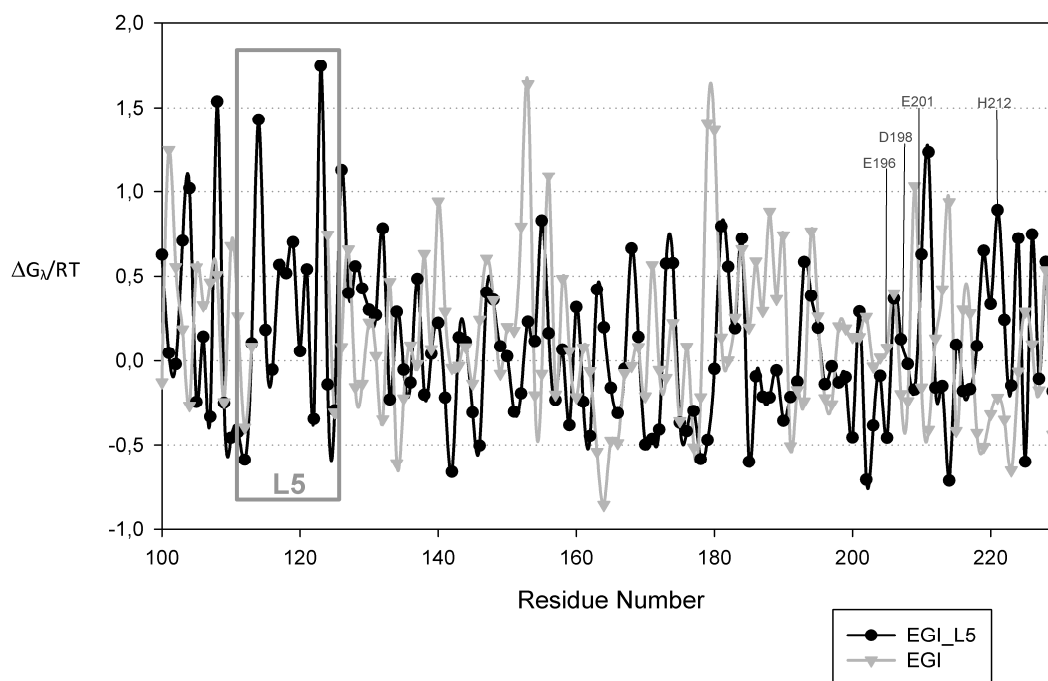


Figure 17: EGI and EGI\_L5 stability coefficients calculated for each residue from molecular mechanics simulations (for simplicity, stability coefficients for residues 100-230 are shown). Stability coefficient  $> 0$  indicates a less rigid structure whereas stability coefficient  $< 0$  indicates a more rigid structure.

#### 4.1.2. Production of Recombinant Enzymes and Fermentation

*egl1\_L5* gene is obtained by overlap PCR extension method (Figure 18). After the loop insertion via overlap PCR extension method, the gene was inserted into pPiczaA plasmid and then transformed into *P. pastoris*. Best clones expressing the recombinant enzymes EGI and EGI\_L5 were selected on BMM-agar plates containing Azo-CMC as the substrate. EGI clone E12 was utilized for fermentation due to its higher activity against Azo-CMC on BMM-agar plates. EGI\_L5 clones were found to exhibit more or less the same activity. As a result, EGI\_L5 clone D5 was chosen for fermentation and further analysis.

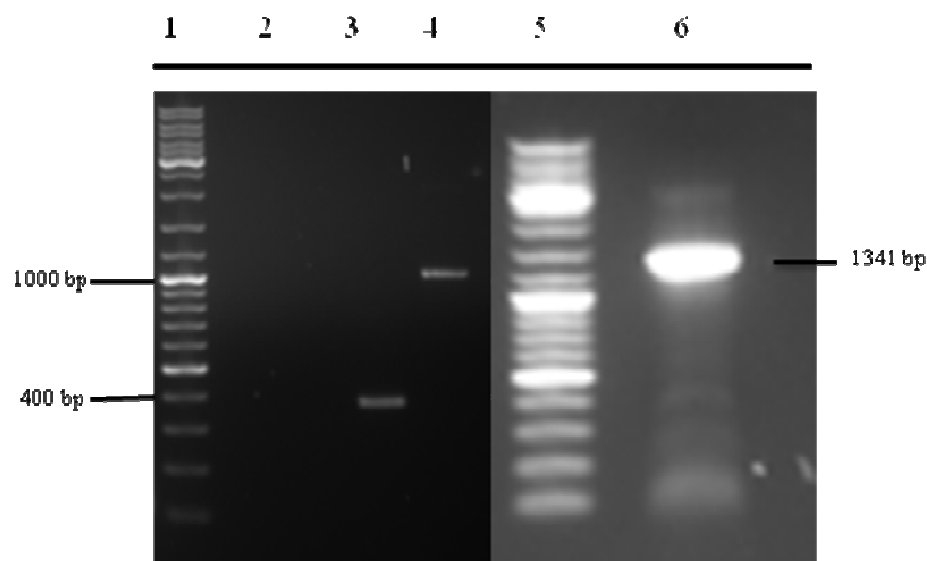


Figure 18: Overlap extension PCR results for the production of *eglL5* gene. Primers in Table 2 are used to produce parts of *eglL* gene. 1 and 5: GeneRuler DNA Ladder Mix (Fermentas), 2: empty well, 3: *eglL5\_1* (~380 bp with overlap extension sequences), 4: *eglL5\_2* (~1000 bp with overlap extension sequences), 6: *eglL5* gene (~1341 bp).

Both EGI clone E12 and EGI<sub>L5</sub> clone D5 were subjected to fed-batch fermentation. Fermentation products were collected at different time points throughout the fermentation process. SDS-PAGE and activity analysis of fermentation products collected at different time points are shown in Figure 19. Growth rates of the clones were followed by measuring and calculating cell dry weights (CDW) for each sample collected at different time points. For each sample, activity against 4-MUC was evaluated according to the rate of formation of 4-MU per minute at 25 °C. Methanol concentration at each time point of fermentation was monitored using a specific methanol probe inside the fermenter. Fermentation data analysis for EGI and EGI<sub>L5</sub> producing clones are shown in Figure 20 and Figure 21, respectively. Growth rates and expression profiles were found to be very similar for EGI and EGI<sub>L5</sub> during fed-batch fermentations. Enzyme activity and enzyme production was found to be increased with the increase in methanol concentration.

There was no enzyme production in glycerol batch and glycerol fed-batch phases in both fermentations as expected. EGI started to be produced after 24 hours and EGI<sub>L5</sub> after 20 hours with methanol induction of the AOX promoter. During methanol fed-batch phase cell growth was minimal. The activities of each recombinant enzyme produced were maximal at the end of the fermentation.

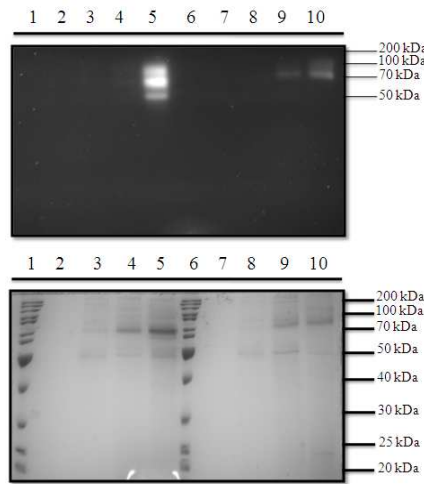


Figure 19. Zymogram (upper picture) and SDS-PAGE (lower picture) analysis of fermentation products. 1-6: PageRuler Protein Ladder Mix (Fermentas), 2: EGI 0h, 3: EGI 16h, 4: EGI 41h, 5: EGI 46h, 7: EGI\_L5 0h, 8: EGI\_L5 14h, 9: EGI\_L5 38h, 10: EGI\_L5 63h.

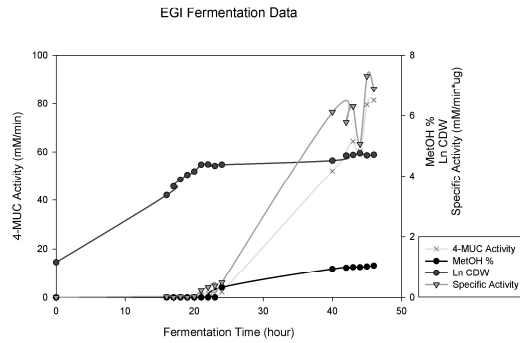


Figure 20: Fermentation data analysis for EGI fed-batch fermentation. 0-16h glycerol batch phase, 16h-20h glycerol fed-batch phase, 20h-46h methanol fed batch phase. CDW: Cell Dry Weight. Specific activity (([4MU] produced per min) / $\mu$ g of produced protein (mM/min\*  $\mu$ g)).

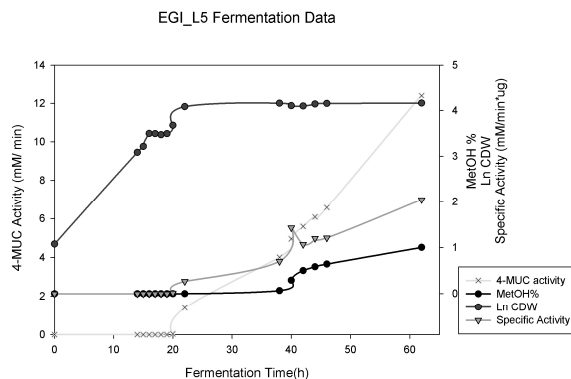


Figure 21: Fermentation data analysis for EGI\_L5 fed-batch fermentation. 0-15h glycerol batch phase, 15h-20h glycerol fed-batch phase, 20h-63h methanol fed batch phase. CDW: Cell Dry Weight. Specific activity (([4MU] produced per min) / $\mu$ g of produced protein (mM/min\*  $\mu$ g)).

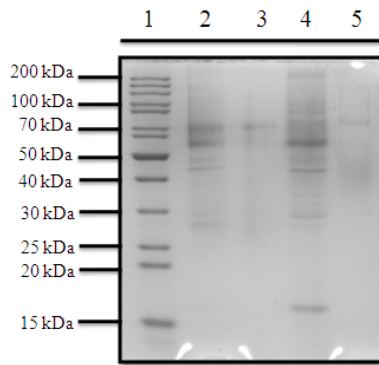


Figure 22: SDS-PAGE of affinity batch purified EGI and EGI\_L5 (12 % SDS-PAGE gel). 1: PageRuler Protein Ladder Mix(Fermentas), 2: EGI fermentation product, 3: purified EGI, 4: EGI\_L5 fermentation product, 5: purified EGI\_L5.

#### 4.1.1. Purification of Recombinant Proteins

Both enzymes were purified to homogeneity after batch affinity purification with RAC. Only one glycosylation form around 70 kDa was purified with RAC (Figure 22). The purified component was the most active among other glycosylation products for both EGI and EGI\_L5. Activity of both purified enzymes indicated that EGI has the same specific activity as EGI\_L5 against 4-MUC at 45 °C, at the same protein concentration (Figure 23).

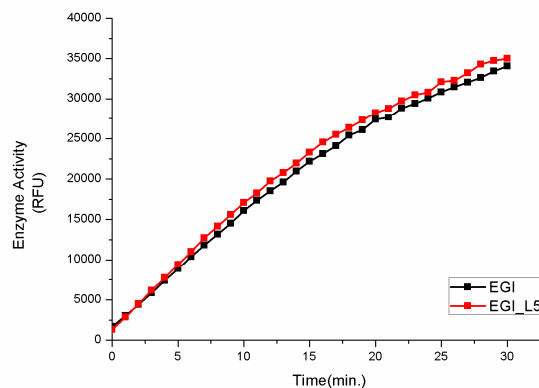


Figure 23: Activity of purified EGI and EGI\_L5 at the same protein concentration against 4-MUC at 45 °C. Activity was determined in terms of relative fluorescence units (RFU) .



#### **4.1.1. Zymogram Analysis**

Activity of the produced recombinant enzymes were monitored qualitatively using a zymogram gel analysis. 4-MUC was used as the substrate for the analysis. EGI produced recombinantly in *Pichia pastoris* was found to have more than one protein band with activity against 4-MUC. These were thought to be different glycosylation (50, 70 kDa) and/or dimerization products (100 kDa) (Figure 19). EGI\_L5 exhibited lower activity against 4-MUC in the zymogram analysis. It was also found to have more than one enzymatically active component, ~100 kDa product being the predominant one, and a probable glycosylation product (~70 kDa). Both forms were found to be active against 4-MUC (Figure 19).

#### **4.1.1. Activity and Stability Analysis**

Temperature activity profiles of EGI and EGI\_L5 produced by fermentation were determined using DNS method and 0.5 % CMC (w/v) as substrate (Figure 24). Activities of both recombinant enzymes at different temperatures exhibited similar profiles. EGI exhibited maximal activity at 45 °C and 55 °C whereas EGI\_L5 was found to be have a maximum activity at 35 °C. Both enzymes showed activity over a broad temperature range (between 15 °C-65 °C). Although EGI\_L5 had a reduced activity at 55 °C with respect to EGI, it showed a slightly higher relative enzyme activity at 65 °C, 75 °C and 85 °C. Both enzymes kept ~40 % of their activities at 75 °C and 30 % of their activities at 85 °C. pH activity profiles of EGI and EGI\_L5 between pH 3 and pH 6 are shown in Figure 25. Both enzymes showed maximal activity at pH 5. pH activity profiles of both enzymes were observed to have followed a very similar pattern. Enzymes were found to be almost inactive at pH 3. Moreover, after 10 minutes incubation at 100 °C EGI and EGI\_L5 were found to retain 94,76 % and 95,41 % of their activities, respectively. The pI's of the EGI and EGI\_L5 calculated with ExPASy (Bjellqvist, et al., 1993) were 4.66 and 5.35, respectively. Although the calculated pI of EGI was shifted almost one pH unit with the introduction of a ten aminoacid loop, its pH profile did not change Both EGI and EGI\_L5 were shown to keep 65 % and 57 % of their activity at 50 °C after 72 hours incubation at 50 °C, respectively. Additionally, it was shown that EGI lost its activity more rapidly upon prolonged incubation at 50 °C (after 24h to 72h) (Figure 26). Moreover, both enzymes have exhibited similar patterns

for their residual enzyme activities upon incubation at 70 °C for 2 hours (Figure 27) . EGI has shown a 5.3 % decrease in residual enzyme activity whereas EGI\_L5 has lost 18 % of its activity after 2 hours of incubation at 70 °C. Kinetic constants for EGI and EGI\_L5 calculated from their activity against 4-MUC at 45 °C were shown in Table 5.

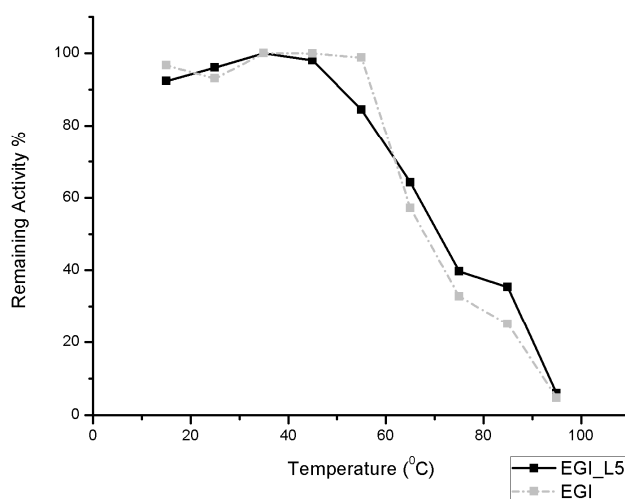


Figure 24: Effect of temperature on hydrolysis of 0.5 % CMC (w/v) by EGI and EGI\_L5. The activity was determined by incubating each enzyme at each temperature for 10 minutes at pH 4.8. The experiments were performed in triplicates with a S.D. of below 10%. The products were analyzed using DNS method for reducing sugars.

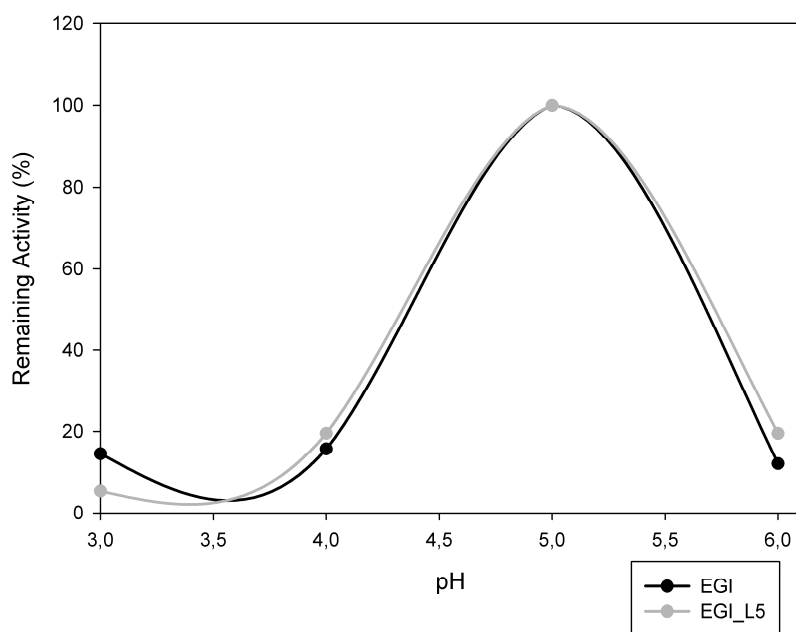


Figure 25: Effect of pH on hydrolysis of 0.5 % CMC (w/v) by EGI and EGI\_L5. The activity was determined by incubating each enzyme at each pH for 10 minutes at 55 °C. The experiments were performed in triplicates with a S.D. of below 10%. The products were analyzed using DNS method for reducing sugars.

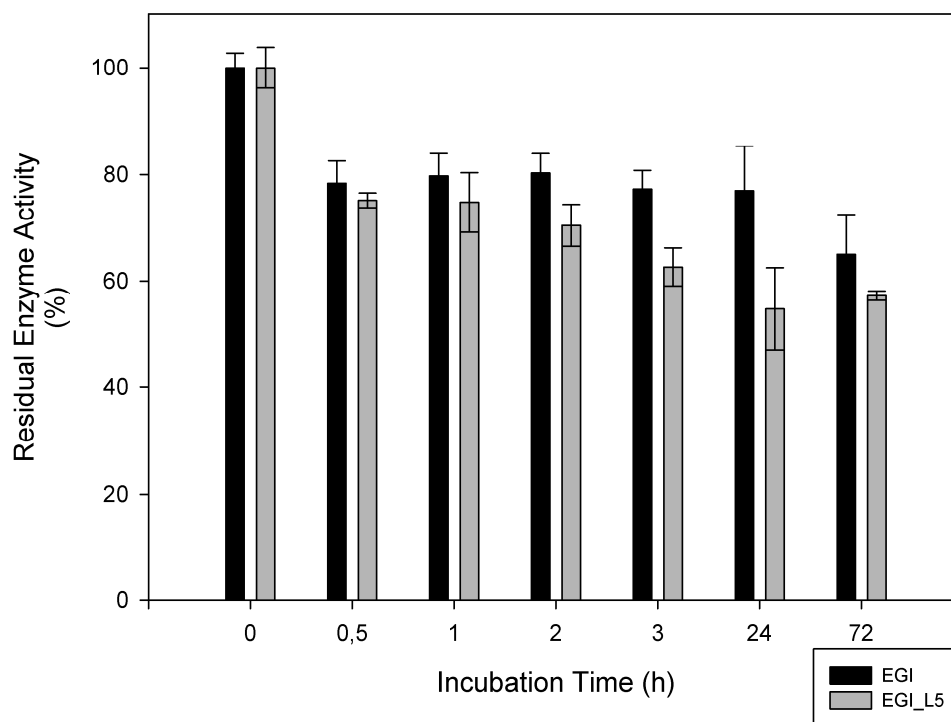


Figure 26: Residual activity of EGI and EGI\_L5 at 50 °C upon incubation for 0 to 72 hours at 50 °C against CMC at pH 4.8.

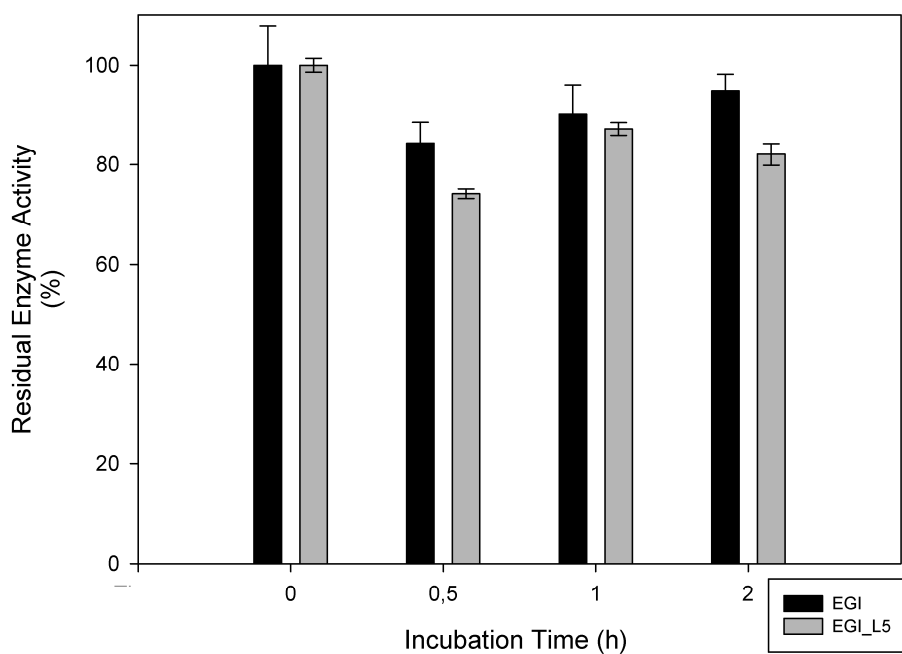


Figure 27: Residual activity of EGI and EGI\_L5 at 50 °C upon incubation for 0 to 2 hours at 70 °C against CMC at pH 4.8.

Table 5: Kinetic constants for EGI and EGI\_L5 calculated from their activity against 4-MUC at 45 °C.

Enzyme	Km(mM)	Kcat(1/sec)	Kcat/Km(1/sec*mM)
EGI	0,332	1,804	5,44
EGI_L5	0,355	0,0257	0,0724

## 4.2. Effect of codon optimization on the production of EGI in *Pichia pastoris*

### 4.2.1. Production of Recombinant Enzymes and Fermentation

Endoglucanase 1 of *Trichoderma reesei* is successfully produced in *Pichia pastoris* for the first time as an active and stable catalyst. *Pichia pastoris* endoglucanase 1 expression yield is comparable to the ones reported in the literature (Nakazawa, et al., 2008; M. E. Penttila, Andre, Saloheimo, Lehtovaara, & Knowles, 1987). The recombinant enzyme exhibited similar activities towards soluble substrates such as CMC and 4-MUC. In addition to endoglucanase 1 expression in *Pichia pastoris*, the effect of codon optimization on the EGI expression was studied.

Comparison of the codon usage of *egl1* and *egl1s* with that of *Pichia pastoris* using a codon usage database (<http://www.kazusa.or.jp/codon>) indicated that frequencies of individual codons of *egl1* is different from those of *Pichia pastoris* (Table 6). Around 81301 codons of *Pichia pastoris* was used by the database to form the *Pichia pastoris* codon usage frequencies. Average GC content of *Pichia pastoris* was calculated to be 42.73 %. Moreover, codon adaptation index (CAI) of *egl1* gene was increased from 0.55 to 0.91 after optimization. Average GC content was reduced to 46 % from 62 %. Four prokaryotic inhibitory motifs and five AT-rich or GC-rich sequence stretches were removed upon optimization.

Colony PCR positive clones were selected for cloning and enzyme production (Figure 28 and 29). Best clones expressing the recombinant enzymes EGI and EGIs were selected on BMM-agar plates containing azo-CMC as the substrate. Clones that exhibited better activity on azo-CMC containing BMM-agar plates were chosen for further analysis. Real-time PCR analysis of AOX promoter regions revealed the copy number of the clones. It was found that EGIs clone E12 and EGI clone C5 have 2.1 and 1.9 copies of *egl1s* and *egl1* expression cassettes on the average, respectively. Even

though we have higher endoglucanase 1 expressing clones, in order to compare the effect of codon optimization on production rates we have selected the clones having the same copy number. As a result, EGIs clone E12 and EGI clone C5 were chosen for fermentation and further analysis.

Table 6: Comparison of codon usage in *Pichia pastoris* genes with that in native and codon-optimized *egl1* genes. Highlighted values show the improved codons.

Aminoacids	Codons	Frequency*			Aminoacids	Codons	Frequency*		
		<i>Pichia pastoris</i>	<i>eg1ls</i>	<i>egl1</i>			<i>Pichia pastoris</i>	<i>eg1ls</i>	<i>egl1</i>
<b>ILE</b>	ATT	0.50	0.33	0.25	<b>SER</b>	TCT	0.29	0.39	0.14
	ATC	0.31	0.67	0.75		TCC	0.20	0.58	0.12
	ATA	0.18	0.00	0.00		TCA	0.18	0.02	0.02
<b>LEU</b>	CTT	0.17	0.04	0.12		TCG	0.09	0.02	0.25
	CTC	0.08	0.00	0.42		AGT	0.15	0.00	0.00
	CTA	0.11	0.00	0.00		AGC	0.09	0.00	0.47
	CTG	0.15	0.00	0.37	<b>TYR</b>	TAT	0.47	0.09	0.17
	TTA	0.16	0.00	0.00		TAC	0.53	0.91	0.83
	TTG	0.33	0.96	0.08	<b>TRP</b>	TGG	1.00	1.00	1.00
<b>VAL</b>	GTT	0.42	1.00	0.05	<b>GLN</b>	CAA	0.61	0.42	0.16
	GTC	0.23	0.00	0.63		CAG	0.39	0.58	0.84
	GTA	0.15	0.00	0.00	<b>ASN</b>	AAT	0.48	0.11	0.03
	GTG	0.19	0.00	0.32		AAC	0.52	0.89	0.97
<b>PHE</b>	TTT	0.54	0.11	0.11	<b>HIS</b>	CAT	0.56	0.00	0.17
	TTC	0.46	0.89	0.89		CAC	0.44	1.00	0.83
<b>MET</b>	ATG	1.00	na**	na**	<b>GLU</b>	GAA	0.56	0.30	0.00
<b>CYS</b>	TGT	0.64	0.95	0.14		GAG	0.44	0.70	1.00
	TGC	0.36	0.05	0.86	<b>ASP</b>	GAT	0.58	0.09	0.14
<b>ALA</b>	GCT	0.45	1.00	0.08		GAC	0.42	0.91	0.86
	GCC	0.26	0.00	0.71	<b>LYS</b>	AAA	0.47	0.00	0.10
	GCA	0.23	0.00	0.04		AAG	0.53	1.00	0.90
	GCG	0.06	0.00	0.17	<b>ARG</b>	CGT	0.16	0.00	0.00
<b>GLY</b>	GGT	0.44	0.86	0.10		CGC	0.05	0.00	0.37
	GGC	0.14	0.00	0.67		CGA	0.10	0.00	0.00
	GGA	0.33	0.14	0.06		CGG	0.05	0.00	0.25
	GGG	0.10	0.00	0.16		AGA	0.48	1.00	0.00
<b>PRO</b>	CCT	0.35	0.04	0.17		AGG	0.16	0.00	0.37
	CCC	0.15	0.00	0.54	<b>STOP</b>	TAA	0.50	na**	na**
	CCA	0.42	0.92	0.04		TGA	0.19	na**	na**
	CCG	0.09	0.04	0.25		TAG	0.31	na**	na**
<b>THR</b>	ACT	0.40	0.96	0.15	*:frequencies of individual codons are shown for each corresponding amino acid. **: not applicable since during cloning pPiczaA vector Start and Stop codons used before a-factor secretion signal and after 6X-His tag and myc epitope, so these sequences are not taken into consideration.				
	ACC	0.26	0.02	0.44					
	ACA	0.24	0.02	0.09					
	ACG	0.11	0.00	0.33					

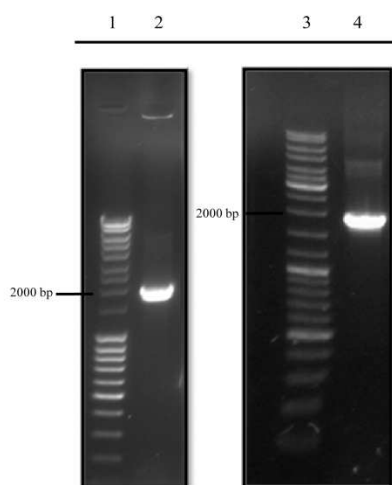


Figure 28: Colony PCR for subcloned *eglls* (lanes 2-4) and *egll* (lanes 6 and 7) genes using 5' and 3' AOX primers. 1: Gene Ruler High Range and Low Range DNA Ladder (mixed, Fermentas), 2: Colony 1(Colony PCR +), 3: Colony 2(Colony PCR -), 4: Colony 3(Colony PCR -), 5: GeneRuler Ladder Mix(Fermentas), 6: Colony 4(Colony PCR +), 7: Colony 5(Colony PCR +).

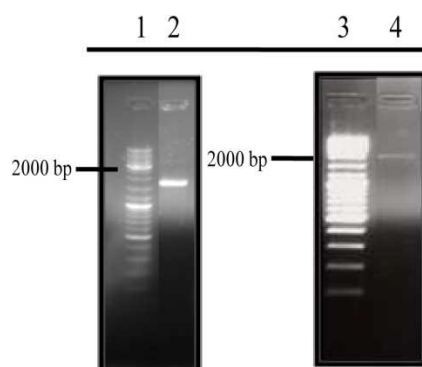


Figure 29: Colony PCR for *eglls* and *egll* expression cassettes transformed into *Pichia pastoris* KM71H using 5' and 3' AOX primers. 1 and 18: GeneRuler Ladder Mix (Fermentas), 2-17: pPiczaA\_eglls transformants (5 is Colony PCR +), 19-21: pPiczaA\_egll transformants (9 is Colony PCR +).

To evaluate enzyme production rates, EGIs clone E12 and EGI clone C5 were grown in shake flasks at 30 °C in BMM. It was found that the average rates of production of enzymes were similar in both clones and wet cell weight calculations indicated that the growth rates of the clones were also similar (Figure 30). It was also observed that due to pH fluctuations, uneven methanol evaporation, and relatively uncontrollable growth conditions, the rates of enzyme production vary with each batch. In order to alleviate this problem and to provide a more controllable, stable growth conditions, these clones were grown in a 5L fermenter. Batch and fed-batch fermentations were applied.

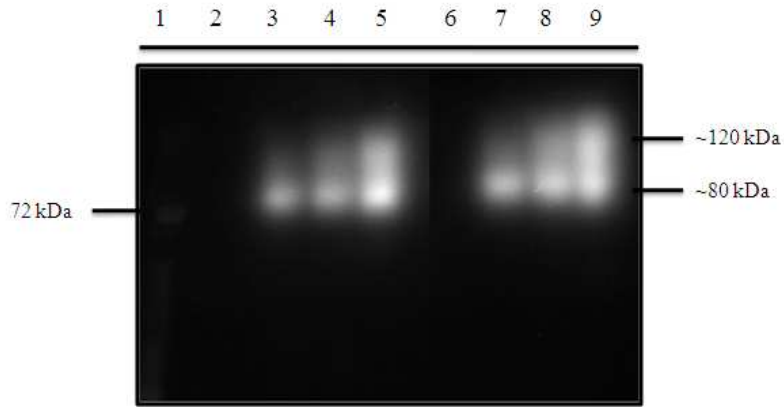


Figure 30: Zymogram analysis of expressed EGIs and EGI in shake flasks against 4-MUC. 1:Prestained Size marker (PageRuler Plus, Fermentas). 2: EGI at 0h, 3: EGI at 24h, 4: EGI at 48h, 5:EGI at 72h, 6: EGIs at 0h, 7: EGIs at 24h, 8: EGIs at 48h, 9: EGIs at 72h.

During batch fermentations both enzymes were found to be produced actively. Zymogram analysis of the culture supernatants revealed that enzymes were active and that their expression levels were similar (Figure 31). Quantitative enzyme assays of the culture supernatants against 4-MUC indicated that EGIs was expressed as a more active protein throughout the batch fermentations than EGI (Figure 32).

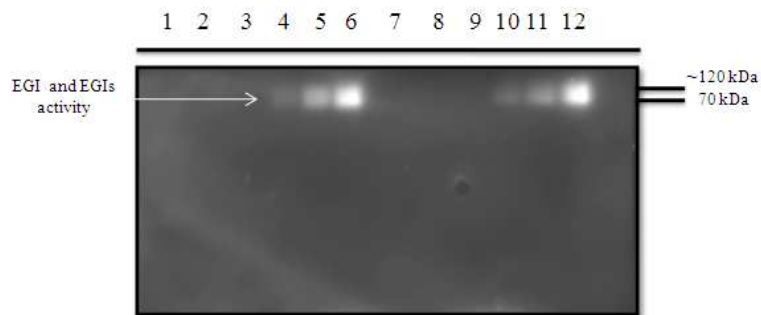


Figure 31: Zymogram analysis of expressed EGIs and EGI as batch fermentation products against 4-MUC. 1-7: Size Marker, PageRuler Ladder Mix (Fermentas). 2: EGI at 0h, 3: EGI at 24h, 4: EGI at 48h, 5: EGI at 72h, 6: EGI at 96h, 8: EGIs at 0h, 9: EGIs at 24h, 10: EGIs at 48h, 11: EGIs at 72h, 12: EGIs at 96h.

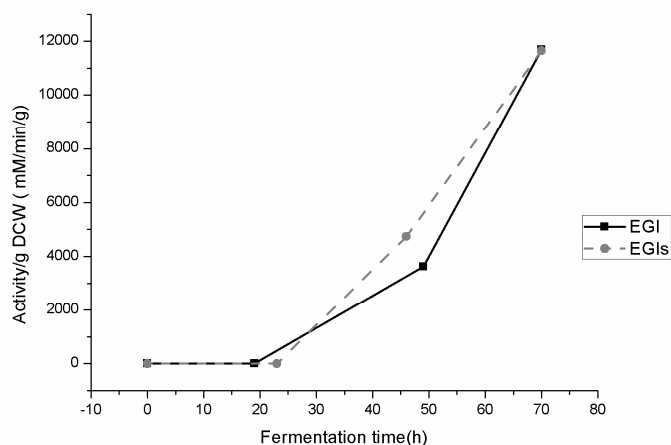


Figure 32: Activities of EGIs and EGI against 4-MUC throughout batch fermentation. Activities were expressed as 4MU produced from 4-MUC per minute divided by cell dry weight. Only samples after methanol induction were shown.

Although total protein produced from batch fermentation of EGIs is less than EGI's, EGIs specific activity against 4-MUC is higher than EGI starting from 30 hours to 60 hours of fermentation time. Both EGIs and EGI exhibited similar specific activities against 4-MUC after 60 hours of fermentation time. Enzyme productivity calculations showed that both EGIs and EGI total protein productivity decreased with time during batch fermentation from the start. EGI total protein productivity decreased to a lesser extent. Despite the fact that both EGIs and EGI fermentation products were degraded during batch production, recombinant EGI and EGIs were resistant to proteolytic degradation by *Pichia* proteases which makes them ideal candidates for enzyme production in *P. pastoris*. EGIs and EGI enzyme activity per g total protein produced were increased to 53.77 RFU/min/g and 43.23 RFU/min/g, respectively. This indicates a 1.24 fold increase in enzyme activity per total protein produced. On the overall, although EGI total protein productivity was found to be more than EGIs on the total protein production level, EGIs was produced as a more active protein product.

Both EGIs clone E12 and EGI clone C5 were subjected to fed-batch fermentation. Fermentation products were collected at different time points throughout the fermentation process. Activity analysis of fermentation products collected at different time points and SDS-PAGE of fermentation products are shown in Figure 33 and 34. Growth rates of the clones were followed by measuring and calculating cell dry weights (CDW) for each sample collected at different time points. For each sample, activity



against 4-MUC was evaluated according to the rate of formation of 4-Methylumbelliferyl (4-MU) per minute at 30 °C. Methanol concentration at each time point of fermentation was monitored using a specific methanol probe inside the fermenter. Growth rates and expression profiles were found to be very similar for EGI and EGIs during fed-batch fermentations. Enzyme activity and enzyme production were found to be increased with the increase in methanol concentration. Although in fed batch fermentations both enzymes have exhibited similar profiles, we can not directly compare their enzyme production rates since their methanol feed rates did not follow the same pattern due to manual adjustments.

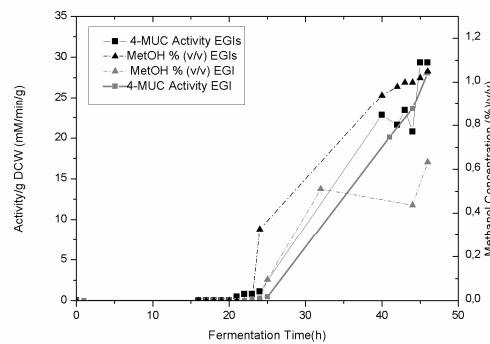


Figure 33: Activities of EGIs and EGI against 4-MUC throughout fed batch fermentation. Methanol concentrations throughout fermentations measured with methanol detection probe were also shown. Glycerol feed rates were kept as 18.15 ml/h/L initial fermentation volume Methanol feed rates were kept between 1-12 ml/h/L fermentation volume. Activities were expressed as 4MU produced per minute divided by cell dry weight.

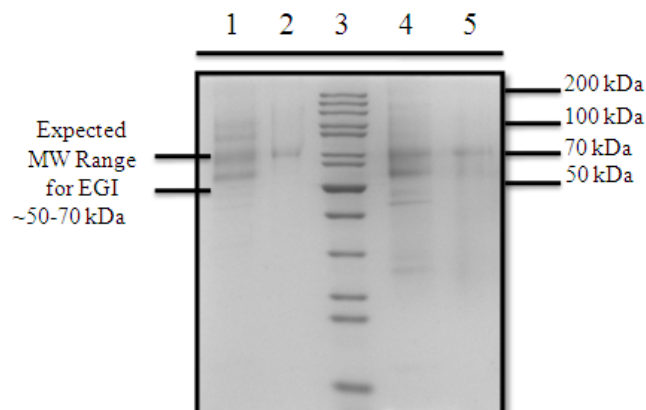


Figure 34 : SDS-PAGE of affinity batch purified EGIs and EGI (12 % SDS-PAGE gel). Expected molecular weight range for EGI and EGIs was 50 to 70 kDa. 1: EGI fermentation product, 2: purified EGI, 3: PageRuler Protein Ladder Mix (Fermentas), 4: EGIs fermentation product, 5: purified EGIs.

### **4.2.2. Purification of Recombinant Proteins**

Both enzymes were purified to homogeneity after one step rapid batch affinity purification with RAC. Only one protein band around 70 kDa was purified with RAC (Figure 34).

### **4.2.3. Activity and Zymogram Analysis**

Activity of the produced recombinant enzymes were monitored qualitatively using a zymogram gel analysis. 4-MUC was used as the substrate for the analysis. EGIs produced recombinantly in *P. pastoris* was found to have more than one protein band with activity against 4-MUC. These were thought to be different glycosylation (50, 70 kDa) and/or dimerization products (100 kDa) (Figure 30 and 31). Expected molecular weight for EGI and EGIs is around 50 to 70 kDa (M. E. Penttila, et al., 1987) which changes according to the glycosylation patterns of the host organism. EGI exhibited lower activity against 4-MUC in the zymogram analysis. It was also found to have more than one enzymatically active component, ~100 kDa product being the predominant one, and a probable glycosylation product around 70 kDa. Both forms were proved to be active against 4-MUC. RAC purified EGI and EGIs were found to exhibit activities of 73.9 and 83.9 RFU/min/ $\mu$ g enzyme against 4-MUC. EGIs was shown to be more active against 4-MUC. Moreover, EGIs was found to exhibit a temperature optimum between 35-55 °C (Figure 35a) and a pH optimum around 5 (Figure 35b) against CMC. Nakazawa et al. cloned EGI catalytic domain in *E. coli* and found that the thermal stability of the recombinant EGI catalytic domain was reduced to less than 80 % after 60 minutes incubation at 50 °C. Our recombinant enzyme was found to retain almost 96 % of its activity upon incubation at pH 4.8 at 50 °C for 5, 24 and 48 hours and 80 % of its activity after 72 hours. This more prolonged stability might be due to the glycosylation of the recombinant enzyme by *P. pastoris*.

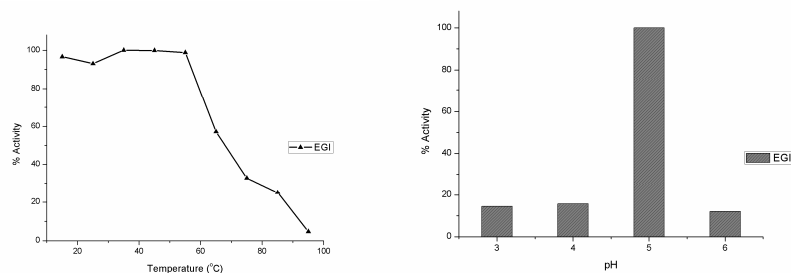


Figure 35: a) Effect of temperature on recombinant EGI activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. b) Effect of pH on recombinant EGI activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. All CMC activity measurements were done as triplicates and all standard deviations were below 10 %.

### 4.3. Cloning and Production of Recombinant Enzymes and Mutants

#### 4.3.1. Production of EGI<sub>BC</sub>

*egl1<sub>bc</sub>* gene is obtained by four rounds of overlap PCR extension and with gel purification of the gene (Figure 36). After the *egl1<sub>BC</sub>* gene production via overlap PCR extension method, the gene was inserted into pPicZαA plasmid and then transformed into *P. pastoris*. Colony PCR was performed on zeocin positive colonies (Figure 37). Best clones expressing the recombinant enzyme EGI<sub>BC</sub> were selected on BMM-agar plates containing Azo-CMC as the substrate. EGI<sub>BC</sub> clone F8 was utilized for fermentation due to its higher activity against Azo-CMC on BMM-agar plates.

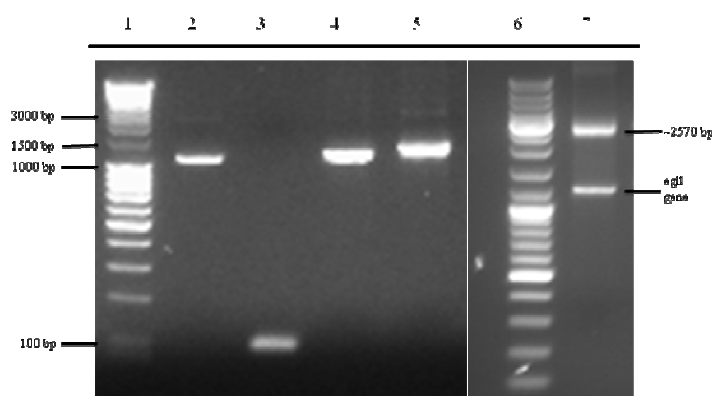


Figure 36: Overlap extension PCR results for the production of *egl1<sub>bc</sub>* gene. Primers in Table 2 are used to produce parts of *egl1* gene. 1: Gene Ruler High Range and Low Range DNA Ladder (mixed, Fermentas), 2: AD; *egl1* gene catalytic domain (~1151 bp with overlap extension sequences), 3: CE: linker region (~88 bp with overlap extension sequences), 4: CD; AD+ CE overlap extension PCR product (~1250 bp with overlap extension sequences), 5: BF; *egl1* gene (~1328 bp with overlap extension sequences), 6: GeneRuler Ladder Mix (Fermentas), 7: GH; *egl1<sub>bc</sub>* gene (~2570 bp).

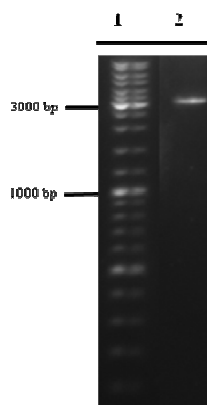


Figure 37: Colony PCR for *eglI\_bc* expression cassette transformed into *Pichia pastoris* KM71H using 5' and 3' AOX primers. 1: GeneRuler Ladder Mix (Fermentas), 2: pPiczaA\_eglI\_bc transformant (Colony PCR +) (~3150 bp→588 bp AOX region+ 2570 bp *eglI\_bc* gene).

EGI\_BC subjected to three different fed-batch fermentations at pH 5, 29 °C; at pH 7, 29 °C and at pH 5, 25 °C. Fermentation products were collected at different time points throughout the fermentation process. Growth rates of the clones were followed by measuring and calculating cell dry weights (CDW) for each sample collected at different time points. For each sample, activity against 4-MUC was evaluated according to the rate of formation of 4-MU per minute per ml at 45 °C. Methanol concentration at each time point of fermentation was monitored using a specific methanol probe inside the fermenter. Fermentation data analysis for EGI\_BC producing clone at under three different conditions are shown in Figure 38, 39 and 40, respectively. Growth rates and expression profiles were found to be very similar for EGI\_BC in comparison to EGI fermentation during fed-batch fermentations at pH 5 and 29 °C. Enzyme activity and enzyme production was found to be increased with the increase in methanol concentration. There was no enzyme production in glycerol batch and glycerol fed-batch phases in both fermentations as expected. EGI\_BC started to be produced after methanol induction of the AOX promoter. During methanol fed-batch phase cell growth was minimal under all three conditions. Fed batch fermentation at pH 7 resulted in a much lower activity compared to pH5, 29 °C and pH5, 25 °C fermentations (Figure 39). Moreover, EGI\_BC enzyme produced at pH5, 25 °C was found to have a lower activity with respect to the enzyme produced at pH5, 29 °C, but a higher activity with respect to the enzyme produced at pH7, 29 °C (Figure 40). Fermentation data analysis have shown an additional decrease in the activity of the enzyme towards the end of the fermentation performed at pH7, 29 °C and pH 5 , 25 °C.

As the zymogram analysis indicates all three fermentation conditions produced degraded enzyme products and highly glycosylated enzymes which are still very active against 4-MUC (Figure 41 and 42). Zymogram analysis has shown that EGI\_BC produced by *P. pastoris* have a 50 kDa component, probably formed as a result of the degradation of the enzyme by *Pichia* proteases, a 85 kDa component and several enzymes that have molecular weights between 100 to 150 kDa, most likely formed as a result of overglycosylation by *Pichia* as the smearing suggests. Only EGI\_BC 50 kDa component could be purified with RAC affinity chromatography (Figure 43). The purified enzyme was found to be active against 4-MUC. EGI\_BC has shown its optimum activity against CMC at pH 5 and 45 °C (Figure 44 a and b).

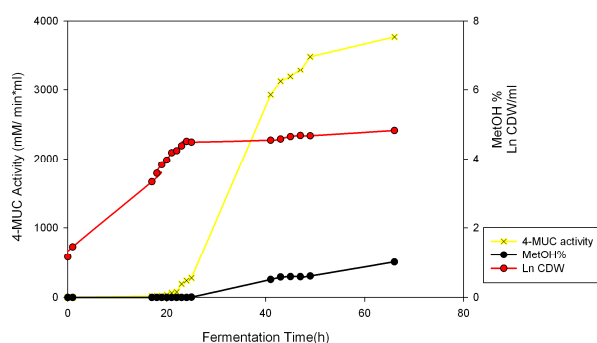


Figure 38: Fermentation data analysis for EGI\_BC fed-batch fermentation (pH 5, 29 °C). 0-17h glycerol batch phase, 17h-25h glycerol fed-batch phase, 25h-65h methanol fed batch phase. CDW: Cell Dry Weight. Activity against 4-MUC is calculated from 4MU (mM) produced per min).

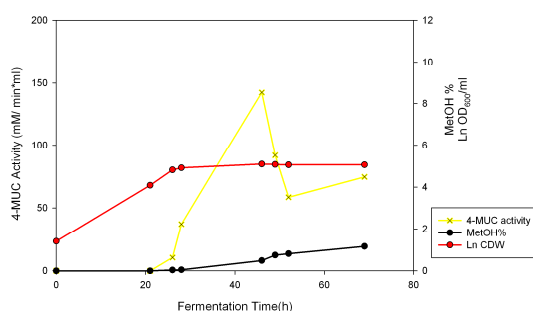


Figure 39: Fermentation data analysis for EGI\_BC fed-batch fermentation (pH 7, 29 °C). 0-21h glycerol batch phase, 21h-26h glycerol fed-batch phase, 26h-69h methanol fed batch phase. CDW: Cell Dry Weight. Activity against 4-MUC is calculated from 4MU (mM) produced per min/ml).

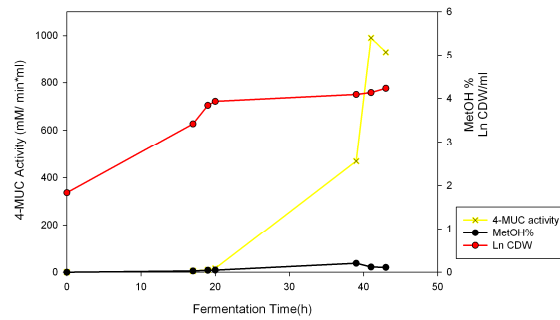


Figure 40: Fermentation data analysis for EGI\_BC fed-batch fermentation (pH 5, 25 °C). 0-17h glycerol batch phase, 17h-20h glycerol fed-batch phase, 20h-43h methanol fed batch phase. CDW: Cell Dry Weight. Activity against 4-MUC is calculated from 4MU (mM) produced per min/ml).

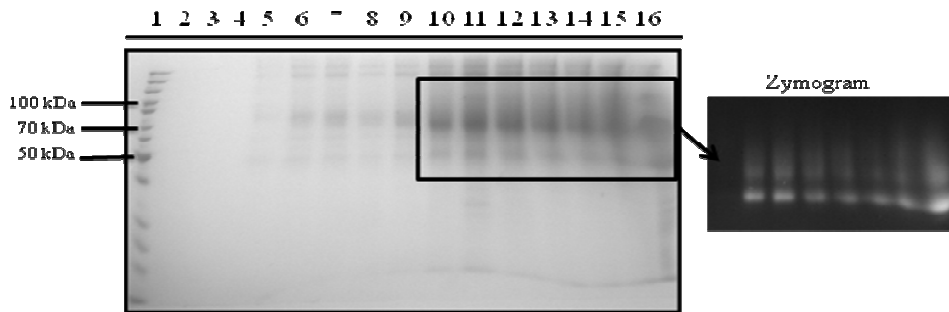


Figure 41: SDS-PAGE and zymogram analysis of expressed EGI\_BC as fed-batch fermentation products against 4-MUC at pH 5, 29 °C. 1: Size Marker, PageRuler Protein Marker (Fermentas). 2: EGI\_BC at 0h, 3: EGI\_BC at 17h, 4: EGI\_BC at 18h, 5: EGI\_BC at 19h, 6: EGI\_BC at 20h, 7: EGI\_BC at 21h, 8: EGI\_BC at 22h, 9: EGI\_BC at 23h, 10: EGI\_BC at 25h, 11: EGI\_BC at 41h, 12: EGI\_BC at 43h, 13: EGI\_BC at 45h, 14: EGI\_BC at 47h, 15: EGI\_BC at 49h, 16: EGI\_BC at 65h.

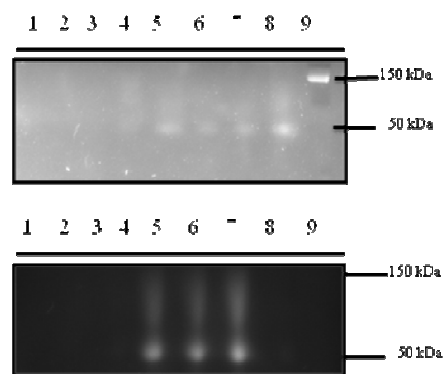


Figure 42: Zymogram analysis of expressed EGI\_BC as fed-batch fermentation products against 4-MUC at pH 7, 29 °C (top picture), at pH 5, 25 °C (bottom picture). 1: EGI\_BC at 0h, 2: EGI\_BC at 21h, 3: EGI\_BC at 26h, 4: EGI\_BC at 28h, 5: EGI\_BC at 46h, 6: EGI\_BC at 49h, 7: EGI\_BC at 52h, 8: EGI\_BC at 69h, 9: Kaledeiscope Prestained Protein Standart (at pH 7, 29 °C); 1: EGI\_BC at 0h, 2: EGI\_BC at 17h, 3: EGI\_BC at 19h, 4: EGI\_BC at 20h, 5: EGI\_BC at 39h, 6: EGI\_BC at 41h, 7: EGI\_BC at 43h, 8: empty well, 9: Size Marker, PageRuler Protein Marker (Fermentas) (at pH 5, 25 °C);

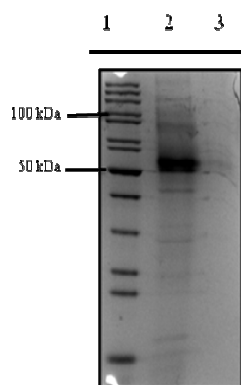


Figure 43 : SDS-PAGE of affinity batch purified EGI\_BC (12 % SDS-PAGE gel). Expected molecular weight range for EGI\_BC was 90 to 120 kDa. 1: PageRuler Protein Ladder Mix (Fermentas), 2: EGI\_BC fermentation product, 3: purified EGI\_BC (~85 kDa and ~50 kDa products).

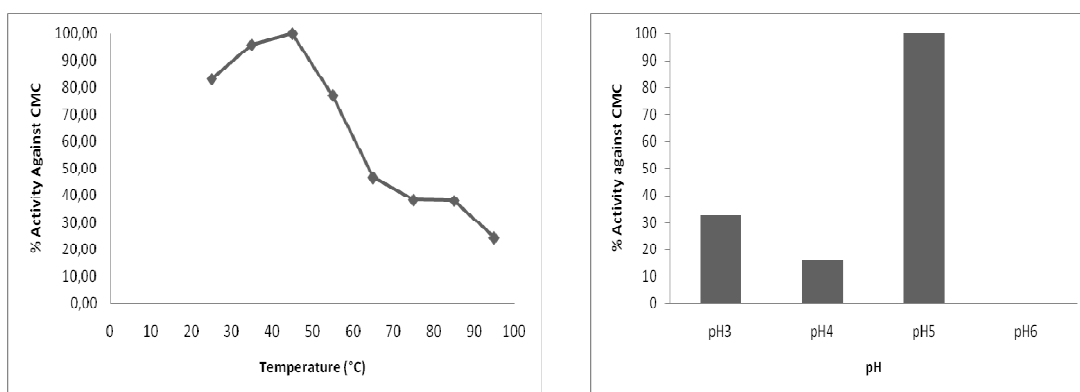


Figure 44: a) Effect of temperature on recombinant EGI\_BC activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. b) Effect of pH on recombinant EGI\_BC activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. All CMC activity measurements were done as triplicates and all standard deviations were below 10 %.

### 4.3.2. Production of EGIII

*egl3* cDNA is obtained using RT-PCR on RNA isolated from *T. reesei* QM9414 strain. After the *egl3* gene production, the gene was inserted into pPiczaA plasmid and then transformed into *P. pastoris*. A plasmid map showing pPiczaA\_egl3 was shown in Figure 45. Colony PCR was performed on zeocin positive colonies (Figure 46). Best clones expressing the recombinant enzyme EGIII were selected on BMM-agar plates containing Azo-CMC as the substrate (Figure 47). EGIII clone C13 was utilized for fermentation due to its higher activity against Azo-CMC on BMM-agar plates and during shake flask experiments.

EGIII was subjected fed-batch fermentation at pH 5, 29 °C. Fermentation products were collected at different time points throughout the fermentation process. Growth rates of the clones were followed by measuring and calculating cell dry weights (CDW) for each sample collected at different time points. For each sample, activity against 4-MUC was evaluated according to the rate of formation of 4-MU per minute per ml at 45 °C. Methanol concentration at each time point of fermentation was monitored using a specific methanol probe inside the fermenter. Fermentation data analysis for EGIII producing clone was shown in Figure 48. Growth rates and expression profiles were found to be very similar for EGIII in comparison to EGI fermentation during fed-batch fermentations at pH 5 and 29 °C. Enzyme activity and enzyme production was found to be increased with the increase in methanol concentration. There was no enzyme production in glycerol batch and glycerol fed-batch phases in both fermentations as expected. EGIII was found to be produced after methanol induction of the AOX promoter. During methanol fed-batch phase cell growth was minimal under all three conditions. The activity of recombinant enzyme against 4-MUC was much lower than EGI activity (Figure 48). The zymogram analysis indicates that the enzyme produced as a single product. No overglycosylation was observed (Figure 49). Zymogram analysis has shown that EGIII produced by *P. pastoris* is a ~27 kDa protein. EGIII enzyme produced could not be purified using RAC as an affinity resin. The enzyme was found to be active against 4-MUC and CMC. EGIII has shown its optimum activity against CMC at pH 5 and 75 °C (Figure 50). After 75 °C, at 85 °C and 95 °C the enzyme activity sharply decreased to 0.



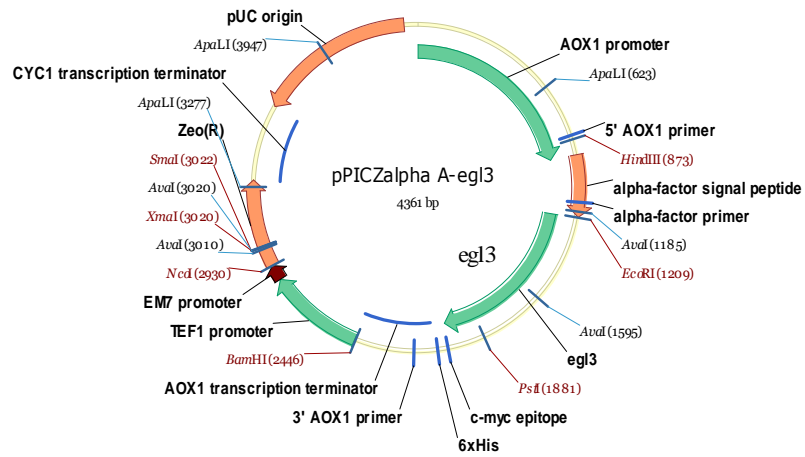


Figure 45: pPic $\alpha$ A-*egl3* plasmid map (drawn with VectorNTI).

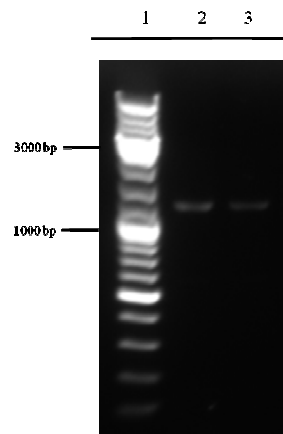


Figure 46: Colony PCR amplification of *pPic $\alpha$ A\_egl3* harboring *P. pastoris* (KM71H) clones with AOX primers. 1: DNA Ladder Mix (Fermentas), 2: Colony PCR (+) clone (~1300 bp), 3: Colony PCR(+) clone (~1300 bp).

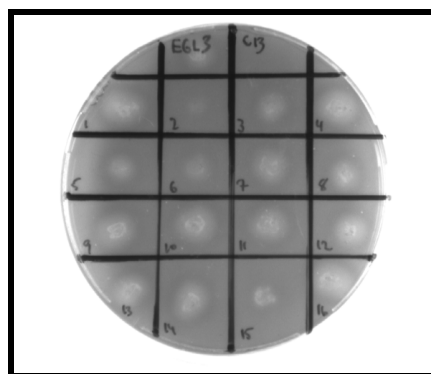


Figure 47: Azo-CMC activity of different *Pichia pastoris* clones producing EGIII. C13: Initial clone, Clones 1-16: 2<sup>nd</sup> round clones (Z series).

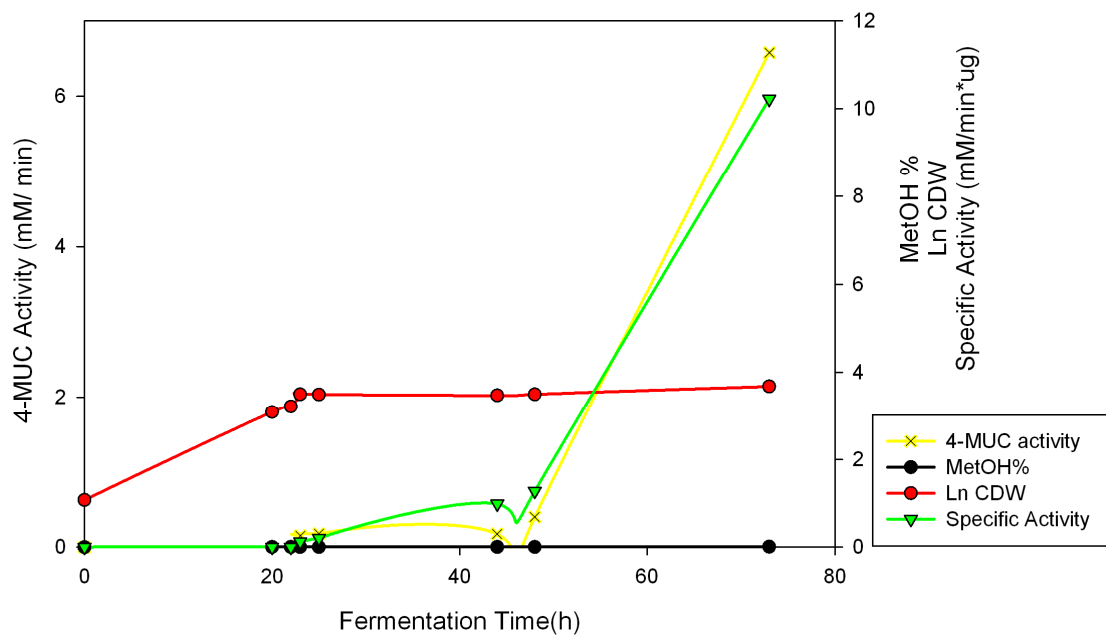


Figure 48: Fermentation data analysis for EGIII clone C13 fed-batch fermentation. 0-20h glycerol batch phase, 20h-23h glycerol fed-batch phase, 23h-73h methanol fed batch phase. CDW: Dry Cell Weight. Specific activity (([4MU] produced per min) / $\mu$ g of produced protein (mM/min\*  $\mu$ g)).

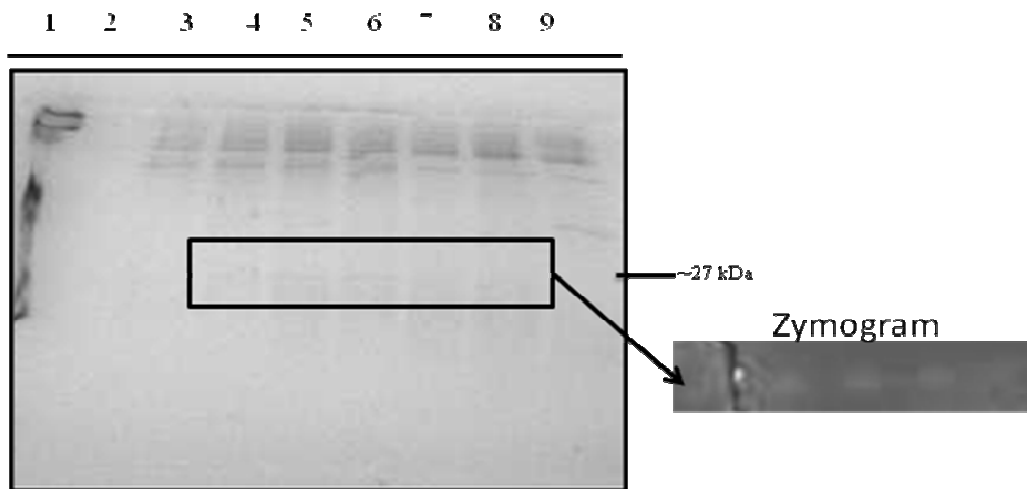


Figure 49: SDS-PAGE and Zymogram analysis of expressed EGIII as fed-batch fermentation products against 4-MUC. 1: Size Marker, Kaledeiscope Prestained Ladder (Biorad). 2: EGIII at 0h, 3: EGIII at 20h, 4: EGIII at 22h, 5: EGIII at 23h, 6: EGIII at 25h, 7: EGIII at 44h, 8: EGIII at 48h, 9: EGIII at 73h.

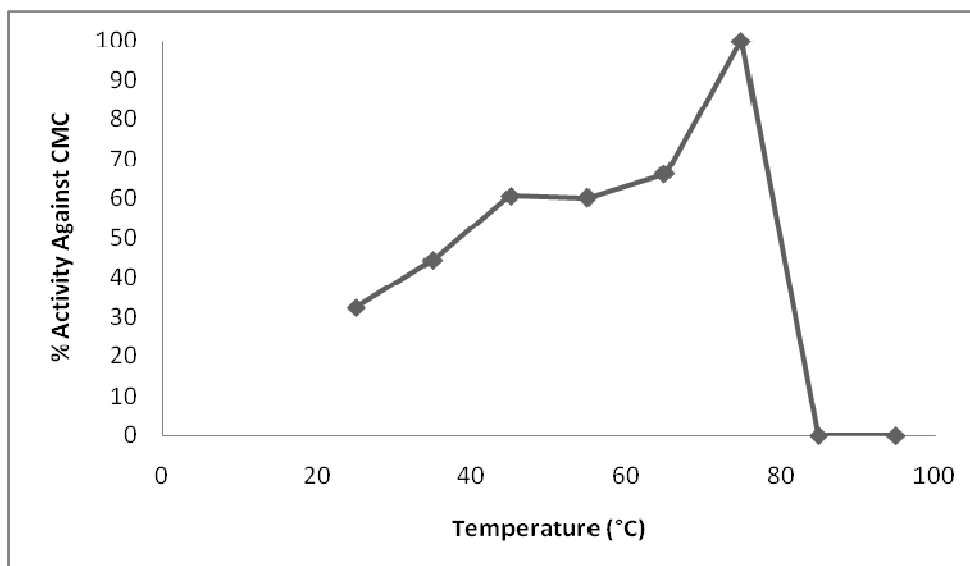


Figure 50: Effect of temperature on recombinant EGI\_BC activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%.

#### 4.3.3. Production of CBHI

*cbh1* cDNA is obtained using RT-PCR on RNA isolated from *T. reesei* QM9414 strain. After the *cbh1* gene production, the gene was inserted into pPiczoB plasmid and then transformed into *P. pastoris*. Colony PCR was performed on zeocin positive colonies (Figure 51). Best clones expressing the recombinant enzyme CBHI were selected according to small scale cultivations in shake flasks. CBHI clones were unable to utilize azo-CMC in BMM-agar plates effectively. CBHI clone X3 was utilized for fermentation due to its higher activity during shake flask experiments.

CBHI was subjected fed-batch fermentation at pH 5, 29 °C. Fermentation products were collected at different time points throughout the fermentation process. Growth rates of the clones were followed by measuring and calculating cell dry weights (CDW) for each sample collected at different time points. For each sample, activity against 4-MUC was evaluated according to the rate of formation of 4-MU per minute per ml at 45 °C. Methanol concentration at each time point of fermentation was monitored using a specific methanol probe inside the fermenter. Growth rates and expression profiles were found to be very similar for CBHI in comparison to EGI fermentation during fed-batch fermentations at pH 5 and 29 °C. Enzyme activity and enzyme production was found to be increased with the increase in methanol concentration. There was no enzyme production in glycerol batch and glycerol fed-batch phases in both fermentations as expected. CBHI was found to be produced after methanol induction of the AOX

promoter. During methanol fed-batch phase cell growth was minimal under all three conditions. The recombinant enzyme did not show any activity against 4-MUC, it has shown activity towards 4-MUL (Figure 52a). But the activity was very low so no activity could be detected in zymogram gels. The SDS-PAGE analysis indicates that the enzyme is not produced as a single product (Figure 52b). SDS-PAGE analysis has shown that CBHI produced by *P. pastoris* is a ~50-70 kDa protein. CBHI enzyme produced could not be purified using RAC as an affinity resin. The enzyme was found to be active against 4-MUL and CMC. CBHI has shown its optimum activity against CMC at pH 6 and at 75 °C (Figure 53). After 75 °C, at 85 °C and 95 °C almost 80 % of the activity was found to be intact.

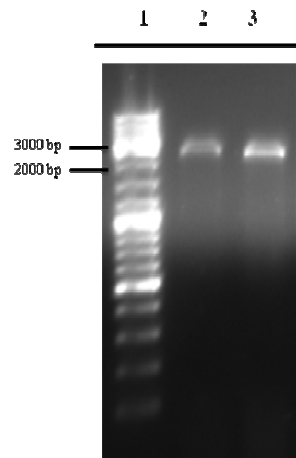


Figure 51: Colony PCR amplification of *pPiczaB\_cbh1* harboring *P. pastoris* (KM71H) clones with AOX primers. 1: DNA Ladder Mix (Fermentas), 2: Colony PCR(+) clone (~2050 bp (1487 bp *cbh1* gene + 558 bp AOX region)), 3: Colony PCR(+) clone (~2050 bp).

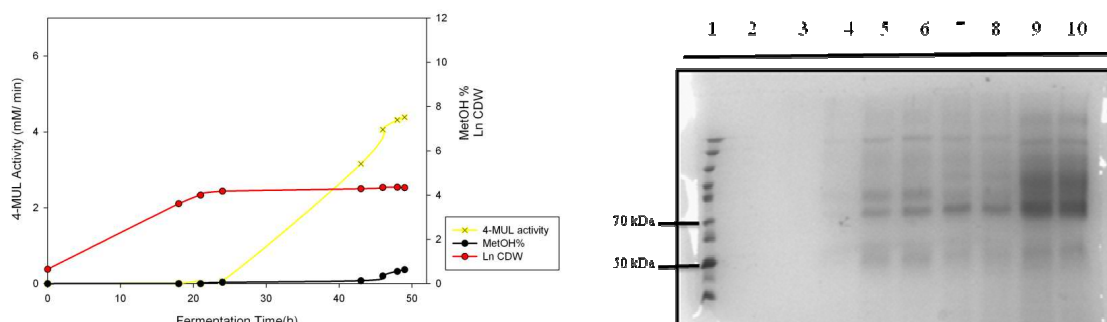


Figure 52: a) Fermentation data analysis for CBHI clone Y2 fed-batch fermentation. 0-20h glycerol batch phase, 20h-24h glycerol fed-batch phase, 24h-49h methanol fed batch phase. CDW: Dry Cell Weight. Specific activity (([4MU] produced per min) /  $\mu$ g of produced protein (mM/min\*  $\mu$ g)). b) SDS-PAGE analysis of expressed CBHI as fed-batch fermentation products against 4-MUC. 1: Size Marker, PageRuler Protein Ladder

Mix (Fermentas), 2: CBHI at 0h, 3: CBHI at 18h, 4: CBHI at 22h, 5: CBHI at 32h, 6: CBHI at 35h, 7: CBHI at 40h, 8: CBHI at 56h, 9: CBHI at 58h, 10: CBHI at 63h.

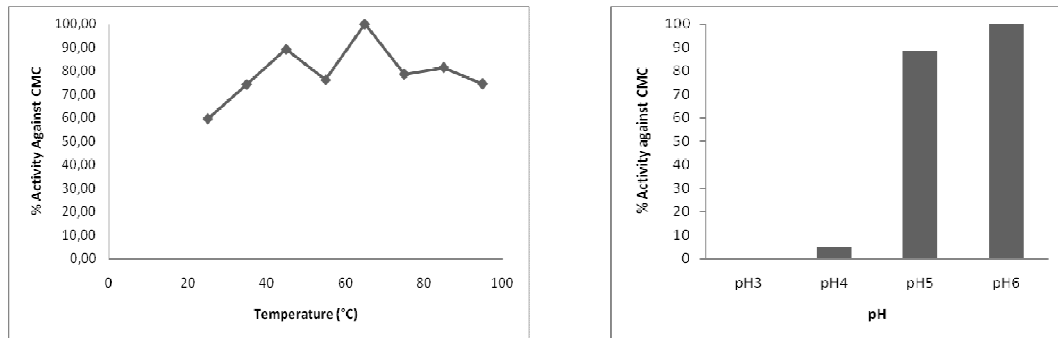


Figure 53: a) Effect of temperature on recombinant CBHI activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. b) Effect of pH on recombinant CBHI activity against CMC. Activity is shown as % activity. Maximum activity was taken as 100%. All CMC activity measurements were done as triplicates and all standard deviations were below 10 %.

#### 4.4. CLEA

Crosslinked enzyme aggregates were prepared from commercial cellulase solution GEMPIL 4L containing endoglucanases, cellobiohydrolases and most probably beta-glucosidases of *T. reesei*. *T. reesei* cellulases are known to exhibit their activity around acidic pH ranges, around pH 5. In all experiments, activity screenings were performed at pH 5 and 55 °C. Effect the precipitant solution on enzyme activity was analyzed by preparing aggregates of Gempil 4L in t-butanol, acetone and ammonium sulfate. enzymes precipitated with t-butanol and acetone have shown similar activities and enzyme precipitated with ammonium sulfate has shown lower activity. Acetone was selected as the most suitable precipitant due to its unharmed effect on the activity of the enzyme to be modified, the cheaper pricing and availability in comparison to t-butanol. It was found that crosslinking pH does not have a great impact on the activity of the of the enzyme (Figure 54). At pH 4 and 6, precipitated Gempil 4L particles exhibited almost the same activity and the activity decreased by 20 % at pH8. Since enzyme activity is not affected much by pH and at lower pH values glutaraldehyde oligomerization might be affected, pH 7.3 was selected for CLEA preparation (Sheldon et al. used pH 7.3 for CLEA preparation).

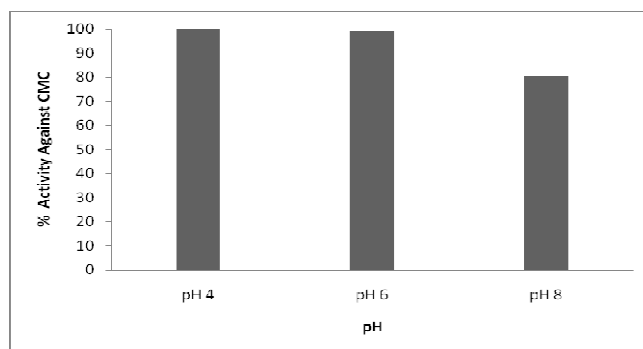


Figure 54: Effect of pH on Gempil 4L activity after precipitation.

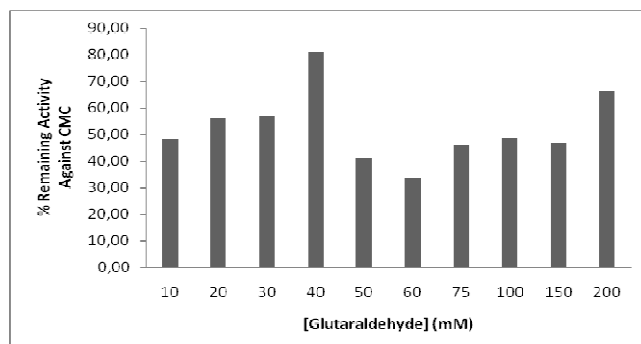


Figure 55: Effect of glutaraldehyde concentration on CLEA activity against CMC after crosslinking.

CLEA activity did not show a decreasing trend with the increase in glutaraldehyde concentration (Figure 55). Rather, it has shown a mixed trend most probably due to small scale preparation problems. Almost 80 % of the activity retained after CLEA preparation with 40 mM glutaraldehyde. This value was the highest activity reached among all glutaraldehyde concentrations. After Gempil 4L CLEA preparation size fractionation of the CLEA particles was performed to evaluate the impact of CLEA size on enzyme activity (Figure 56). After fractionation, Gempil 4L CLEA with 7 different sizes was obtained. These fractions were: B: 25 $\mu$ m >CLEA size>10 $\mu$ m, C: 33 $\mu$ m >CLEA size>25 $\mu$ m, D: 45 $\mu$ m >CLEA size>33 $\mu$ m, E: 77 $\mu$ m >CLEA size>45 $\mu$ m, F: 154 $\mu$ m >CLEA size>77 $\mu$ m, G: 288 $\mu$ m >CLEA size>154 $\mu$ m, H: 1980 $\mu$ m >CLEA size>288 $\mu$ m. Evaluation of the activity of these particles against CMC has indicated that all fractions had exhibited almost the same activity against CMC (Figure 56). But there was a slight increase in the activity of larger sized fractions, E, F and G in comparison to B, C and D.

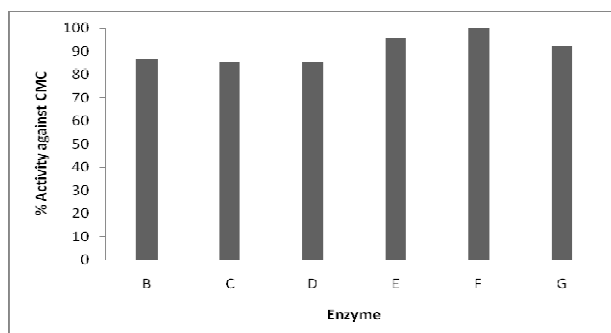


Figure 56: Effect of CLEA size on CLEA activity. CLEA sizes: Gempil 4L: ~10 nm, B: 25 $\mu$ m >CLEA size>10 $\mu$ m, C: 33 $\mu$ m >CLEA size>25 $\mu$ m, D: 45 $\mu$ m >CLEA size>33 $\mu$ m, E: 77 $\mu$ m >CLEA size>45 $\mu$ m, F: 154 $\mu$ m >CLEA size>77 $\mu$ m, G: 288 $\mu$ m >CLEA size>154 $\mu$ m, H: 1980 $\mu$ m >CLEA size>288 $\mu$ m.

CLEA particles prepared from recombinant enzymes were optimized according to CLEA prepared from Gempil 4L but since the enzyme concentrations were lower for recombinant enzymes, lower glutaraldehyde concentrations were used. Crosslinking pH was kept at pH 7.3. EGI CLEA kept 65 % of its activity at mM glutaraldehyde concentration and EGI\_L5 CLEA kept almost 50 % of its activity intact at 0,33 mM Glutaraldehyde concentration. At higher glutaraldehyde concentrations EGI\_L5 CLEA activity against 4-MUC has rapidly diminished (Figure 57). EGI CLEA and EGI\_L5 CLEA particles have exhibited a catalytic activity of 2.2 CMC units/ml and 2.08 CMC units/ml per mg CLEA particle after crosslinking (Figure 58).

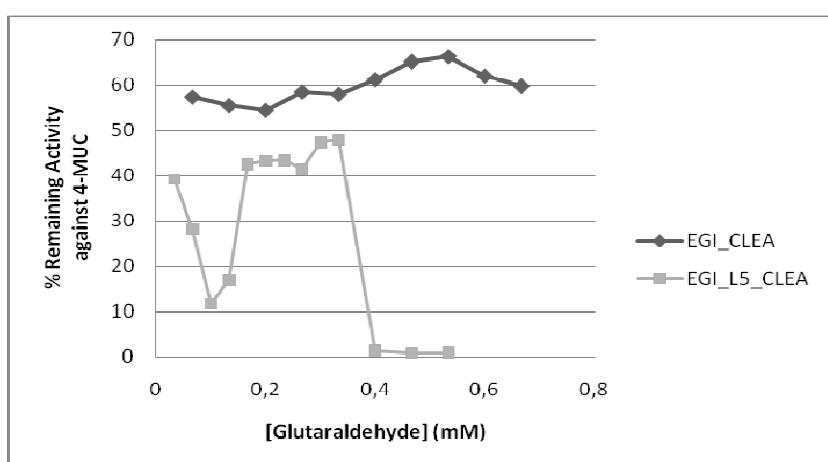


Figure 57: Effect of glutaraldehyde concentration on CLEA prepared from EGI and EGI\_L5.

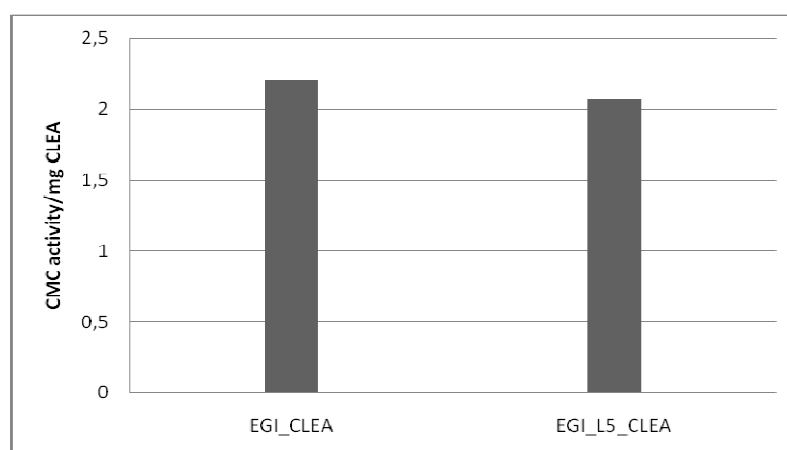


Figure 58: Effect of crosslinking on enzyme activity at pH 5, 55°C.



#### **4.5. Fabric Tests**

Effect of crosslinked and recombinant enzymes on viscose biopolishing was evaluated by analyzing pilling values and fabric bursting strength after enzyme application. Effect of enzyme dosage on pilling and fabric strength was analyzed by applying different amounts of Gempil 4L (G) and Gempil 4L CLEA (C) particles on the fabrics. It was found that for the first trials, Gempil 4L dosage has not improved the pilling rates but fabric strength was reduced with the increase in the enzyme amount. On the other hand CLEA application improved the pilling notes when applied in 1X and 2X concentrations (G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric, 2X is two times the necessary amount/g fabric and so on) (Table 7). The bursting strength of the fabrics was found to be optimal when 1X concentration of the enzyme is applied for 50 minutes or 2 hours. The increase in enzyme dosage of Gempil 4L CLEA decreased the bursting strength values of the viscose knitted fabrics to a much lesser extent at all enzyme concentrations than Gempil 4L. The biopolishing experiments were repeated for 1X and 2X enzyme dosage and effect of process times were evaluated for Gempil 4L CLEA (Table 8). It was shown that best pilling result was obtained by application of 2X CLEA for 2 hours and CLEA particles without fractionation results in nonhomogenous distribution of the enzymes on the fabric and uneven biopolishing results that cannot be repeated. Application of recombinant enzymes EGI, EGI\_L5 and EGI\_BC improved the fabric pilling notes and strength of the fabric has not reduced by the action of the enzymes (Table 8). EGIII did not seem to exhibit any effect on fabric pilling rates and bursting strengths. Application of EGI\_BC in combination with Gempil 4L did not cause any improvements in the pilling rates, instead it has caused a decrease in the fabric strength. Preliminary experiments with CLEA of EGI and EGI\_L5 have shown that their application did not improve the pilling notes but prevented tensile strength loss.

Impact of the recombinant enzymes, Gempil 4L and Gempil 4L CLEA on fabric appearance was evaluated by taking digital photographs after biopolishing (Figure 59 and 60). The effect of all enzymes on the fabric surfaces could not be distinguished with visual inspection. The microfibrils that cause pill formation were present in all fabrics. CLEA particles remained on the surface of the enzyme during biopolishing as indicated by Figure 60.

Table 7: Pilling test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) or recombinant enzymes. G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric, 2X is two times the necessary amount/g fabric and so on.

Sample	Treatment	PILLING					BURSTING STRENGTH		
		125 RPM	500 RPM	1000 RPM	1500 RPM	2000 RPM	Pressure (kPa)	Distensi on (mm)	Time (sec)
<b>G 0.5X</b>	<b>Fabric treated with 0,5X Gempil 4L</b>	1,5	1,5	1,5	1,5	1,5	520,0	13,0	20,0
G 1X	Fabric treated with 1X Gempil 4L for 50 minutes	1,5	1,5	1,5	1,5	1,5	542,0	13,0	21,0
<b>G 1X</b>	<b>Fabric treated with 1X Gempil 4L for 2 hours</b>	1,5	1,5	1,5	1,5	1,5	473,0	12,0	18,0
G 2X	Fabric treated with 2X Gempil 4L for 2 hours	1,5	1,5	1,5	1,5	1,5	454,0	12,0	17,0
<b>G 4X</b>	<b>Fabric treated with 4X Gempil 4L for 2 hours</b>	1,5	1,5	1,5	1,5	1,5	456,0	12,0	17,0
G 8X	Fabric treated with 8X Gempil 4L for 2 hours	1,5	1,5	1,5	1,5	1,5	455,0	13,0	17,0
<b>C 0.5X</b>	<b>Fabric treated with 0,5X CLEA for 2 hours</b>	1,5	1,5	1,5	1,5	1,5	563,0	13,0	22,0
C 1X	Fabric treated with 1X CLEA for 50 minutes	2,5	1,5	1,5	1,5	1,5	611,0	14,0	24,0
<b>C 1X</b>	<b>Fabric treated with 1X CLEA for 2 hours</b>	2	2	1,5	1,5	1,5	581,0	13,0	22,0
C 2X	Fabric treated with 2X CLEA for 2 hours	2	2	2	2	2	472,0	12,0	18,0
<b>C 4X</b>	<b>Fabric treated with 4X CLEA</b>	1,5	1,5	1,5	1,5	1,5	538,0	13,0	21,0
C 8X	Fabric treated with 8X CLEA	1,5	1,5	1,5	1,5	1,5	542,0	13,0	21,0

Table 8: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different treatment times or recombinant enzymes G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric.

Sample	Treatment	PILLING					BURSTING STRENGTH		
		125 RPM	500 RPM	1000 RPM	1500 RPM	2000 RPM	Pressure (kPa)	Distension (mm)	Time (sec)
<b>Fabric</b>	<b>No treatment</b>	1,5	1,5	1,5	1,5	1,5	563	11,2	34,1
Fabric	Buffer treatment	1,5	1,5	1,5	1,5	1,5	559	13,9	34,4
<b>C1X</b>	<b>Fabric treated with 1X CLEA for 30 minutes</b>	1,5	1,5	1,5	1,5	1,5	521	13,5	31,6
C1X	Fabric treated with 1X CLEA for 1 hour	1,5	1,5	1,5	1,5	1,5	500	14,3	30,3
<b>C1X</b>	<b>Fabric treated with 1X CLEA for 2 hours</b>	1,5	1,5	1,5	1,5	1,5	485	13,5	29,4
C2X	Fabric treated with 2X CLEA for 30 minutes	1,5	1,5	1,5	1,5	1,5	526	13,9	32
<b>C2X</b>	<b>Fabric treated with 2X CLEA for 1 hour</b>	1,5	1,5	1,5	1,5	1,5	485	13	30
C2X	Fabric treated with 2X CLEA for 2 hours	2	1,5	1,5	1,5	1,5	475	12,3	28,8
<b>EG1</b>	<b>Treatment with EG1</b>	2,5	1,5	1,5	1,5	1,5	534	13	32,4
EG1_L5	Treatment with EG1_L5	2,5	1,5	1,5	1,5	1,5	525	12,6	31,9
<b>EG1_BC</b>	<b>Treatment with EG1_BC</b>	2	1,5	1,5	1,5	1,5	515	13,7	31,2
EG1_BC	Treatment with EG1	1,5	1,5	1,5	1,5	1,5	545	14	33
<b>EG1_BC +Gempil 4L</b>	<b>Treatment with EG1_BC-Gempil 4L</b>	1,5	1,5	1,5	1,5	1,5	426	13	26

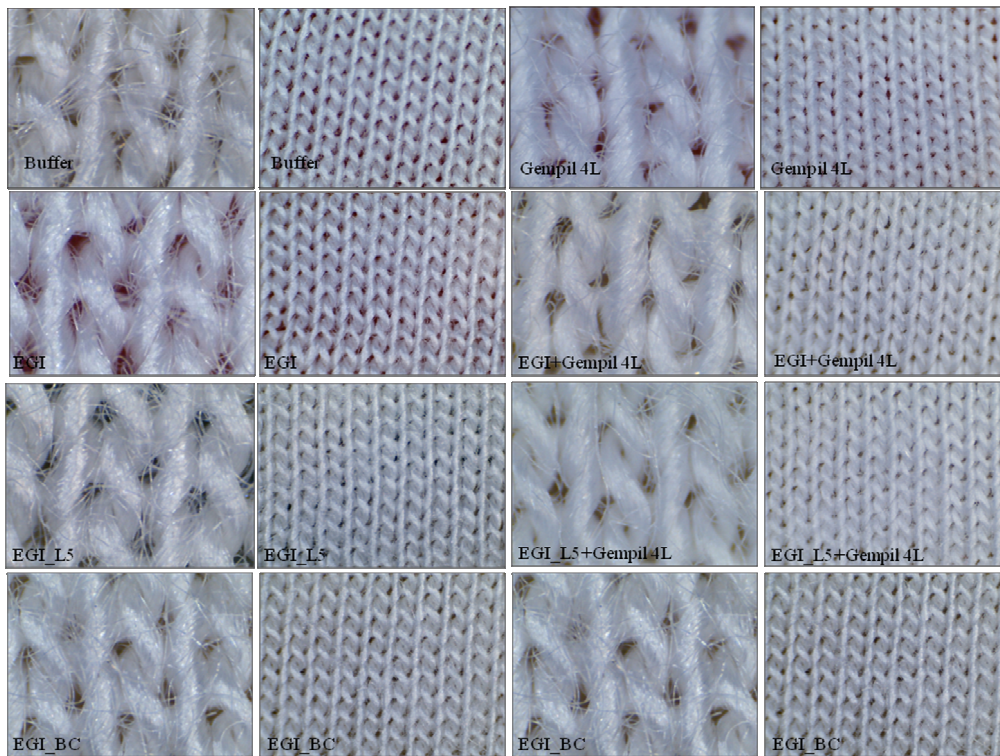


Figure 59: Digital light microscope photographs of viscose knitted fabrics under ~15X and ~400X magnification.

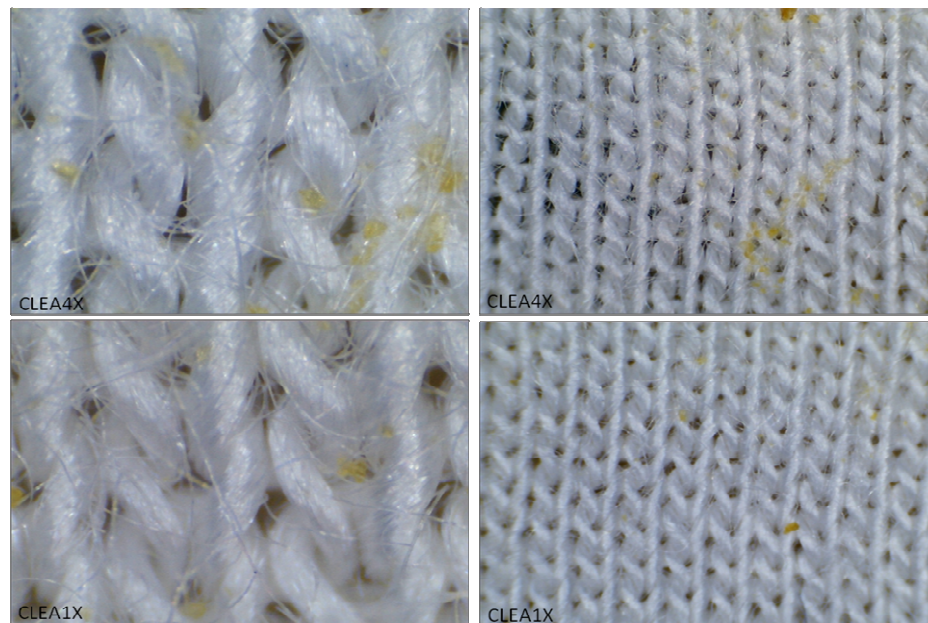


Figure 60: Digital light microscope photographs of viscose knitted fabrics treated with 1X and 4X CLEA under ~15X and ~400X magnification.

Effect of crosslinked and wild type (Gempil 4L) enzymes on viscose woven fabric biopolishing was evaluated by analyzing pilling values and fabric bursting strength after enzyme application. Effect of enzyme dosage on pilling and fabric strength was analyzed by applying different amounts of Gempil 4L (G) and Gempil 4L-CLEA (C) particles on the fabrics. It was found that both enzymes improved the pilling values of the viscose woven fabric from 0.5X to 4X dosage (Table 9). Best effect on pilling was obtained by application of Gempil 4L at 2X dosage and the bursting strength of the fabric did not reduce much with respect to buffer treated fabric (Table 10). Application of Gempil-CLEA at 0.5X dosage has given best results for pilling and the bursting strength value of the fabric was better than Gempil 4L treated fabric (Table 9 and 10).

Table 9: Pilling test results for viscose woven fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C). G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric.

Sample	Treatment	125 RPM	500 RPM	1000 RPM	1500 RPM	2000 RPM
E	Fabric treated with Buffer	1,5	1,5	1,5	1,5	1,5
A:G1X	Fabric treated with 1X Gempil4L	3,5	3,5	3,5	3,5	3,5
F:G2X	Fabric treated with 2X Gempil4L	3	3	3,5	4	4,5
G:G4X	Fabric treated with 4X Gempil4L	4,5	4	3,5	3,5	3,5
H:C0.5X	Fabric treated with 0,5X CLEA	2	3,5	4	4,5	4,5
B:C1X	Fabric treated with 1X CLEA	4	3,5	3,5	3	3
C:C2X	Fabric treated with 2X CLEA	3	2,5	2,5	2,5	2,5
D:C4X	Fabric treated with 4X CLEA	3,5	3,5	3,5	3,5	3,5
I:C8X	Fabric treated with 8X CLEA	1,5	1,5	1,5	1,5	1,5
J:C16X	Fabric treated with 16X CLEA	1,5	1,5	1,5	1,5	1,5

Table 10: Bursting strength test results for viscose woven fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C).

Sample	Treatment	Bursting Strength(kPa)	Distension (mm)	Time (sec)	Note
E	Fabric treated with Buffer	656,0	7,1	20,7	not burst
H	Fabric treated with 0.5X CLEA	648,0	7,2	20,7	not burst
F	Fabric treated with 2X Gempil 4L	614,0	9,4	15,5	burst

After size fractionation of CLEA particles, effect of CLEA size on viscose knitted fabric pilling and strength was evaluated. Experiments were performed in two different apparatus: in a hybridization chamber, a liquor ratio of 1:9 was used with constant rotation; in a special plastic container with a embroidery hoop like apparatus. This apparatus acted as a holder for the fabric. A liquor ratio of 1:50 was used with the apparatus and the container housing the apparatus was put in a shaker incubator for biopolishing under constant shaking. CLEA preparations from two batches were used for biopolishing in hybridization chamber (B, C, D, E, G, H from first batch, F2, F3, F4, F8 from second batch, CLEA sizes were shown in Table 11). Application of B, C, D and E from batch one improved the pilling notes, best effect was obtained by application of fraction E, bursting strength value was better than Gempil 4L and all other fractions from batch one (Table 11). Application of F2, F3, F4 and F8 from batch two, improved the pilling notes, best effect was obtained by application of fraction F8 in terms of pilling and fabric strength, bursting strength value was better than Gempil 4L and all other fractions from batch one and two (Table 11). Application of fractions F3, F4, F8 and C in special apparatus for biopolishing did not result in any improvements in pilling notes or strength of the fabrics (Table 12).

Table 11: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different sizes in hybridization chamber G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric.

Sample	Treatment	PILLING					BURSTING STRENGTH		
		125 RPM	500 RPM	1000 RPM	1500 RPM	2000 RPM	Pressure (kPa)	Distension (mm)	Time (sec)
<b>Fabric</b>	<b>No treatment</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>563</b>	<b>11,2</b>	<b>34,1</b>
<b>Fabric</b>	<b>Buffer treatment</b>	1,5	1,5	1,5	1,5	1,5	559	13,9	34,4
<b>Gempil 4L</b>	<b>Gempil 4L treatment</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>425</b>	<b>12</b>	<b>26</b>
<b>CLEA_B</b>	<b>25µm &gt;CLEA size&gt;10µm</b>	2,5	1,5	1,5	1,5	1,5	444	14	33
<b>CLEA_C</b>	<b>33µm &gt;CLEA size&gt;25µm</b>	<b>2,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>432</b>	<b>13</b>	<b>33</b>
<b>CLEA_D</b>	<b>45µm &gt;CLEA size&gt;33µm</b>	2,5	1,5	1,5	1,5	1,5	448	15	34
<b>CLEA_E</b>	<b>77µm &gt;CLEA size&gt;45µm</b>	<b>2,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>455</b>	<b>13</b>	<b>35</b>
<b>CLEA_G</b>	<b>288µm &gt;CLEA size&gt;154µm</b>	1,5	1,5	1,5	1,5	1,5	531	13	32
<b>CLEA_H</b>	<b>1980µm &gt;CLEA size&gt;288µm</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>503</b>	<b>13</b>	<b>31</b>
<b>CLEA_F2</b>	<b>1980µm &gt;CLEA size&gt;288µm</b>	2,5	1,5	1,5	1,5	1,5	421	14	33
<b>CLEA_F3</b>	<b>288µm &gt;CLEA size&gt;154µm</b>	<b>2,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>425</b>	<b>14</b>	<b>33</b>
<b>CLEA_F4</b>	<b>154µm &gt;CLEA size&gt;77µm</b>	2,5	1,5	1,5	1,5	1,5	410	13	32
<b>CLEA_F8</b>	<b>25µm &gt;CLEA size&gt;10µm</b>	<b>2,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>1,5</b>	<b>471</b>	<b>15</b>	<b>36</b>

Table 12: Pilling and bursting strength test results for viscose knitted fabrics treated with Gempil 4L (G) and Gempil 4L-CLEA (C) with different sizes in special apparatus G1X=180 µl Gempil 4L/ g fabric. C1X= mg CLEA equivalent to 180 µl Gempil 4L/ g fabric.

Sample	Treatment	PILLING					BURSTINGSTRENGTH		
		125 RPM	500 RPM	1000 RPM	1500 RPM	2000 RPM	Pressure (kPa)	Distension (mm)	Time (sec)
<b>Fabric</b>	Notreatment	1,5	1,5	1,5	1,5	1,5	563	11,2	34,1
<b>CLEA_F3</b>	288µm >CLEA size>154µm	1,5	1,5	1,5	1,5	1,5	544	11	33
<b>CLEA_F4</b>	154µm >CLEA size>77µm	1,5	1,5	1,5	1,5	1,5	437	9	32
<b>CLEA_F8</b>	25µm >CLEA size>10µm	1,5	1,5	1,5	1,5	1,5	459	9	34
<b>CLEA_C</b>	33µm >CLEA size>25µm	1,5	1,5	1,5	1,5	1,5	391	10	30

## CHAPTER 5

### 5. DISCUSSION

#### *5.1. Modelling and production of Trichoderma reesei endoglucanase 1 and its mutant in Pichia pastoris*

All the proposed loop insertions were placed away from the active site in order not to interfere with enzyme activity since enzyme active site is located on an open cleft. MD and MM simulations showed that among all loop inserted structures, EGI\_L5 was the only structure that behaved similar to the native structure EGI in terms of RMSD during the simulations. EGI\_L5 exhibited better properties than all other loop mutants in the last 2000 ps of the simulations. MM simulation supported the results of MD simulations. Insertion of a ten aminoacid loop between 112<sup>th</sup> and 113<sup>th</sup> residues of EGI, on the opposite side of the active site, did not decrease the overall stability of EGI. Active site of the EGI became more rigid, thus more stable with the addition of a flexible loop.

*Trichoderma reesei* has been known to produce many endoglucanases, EGI being a major one. EGI comprises 5-10 % of the proteins secreted by *Trichoderma reesei* (M. Penttila, Nevalainen, Ratto, Salminen, & Knowles, 1987). *Trichoderma reesei* cellulase system has been known to exhibit its maximum activity around pH 5 (Bommarius, et al., 2008) and around 50 °C to 60 °C (our groups unpublished results). Our expression results of EGI and EGI\_L5 in *Pichia pastoris* were consistent with these data. Heterologous expression of the codon optimized *eglI* gene and its mutant in *Pichia pastoris* produced glycosylated enzyme products. Several studies indicate a heterogenous glycosylation of *Trichoderma reesei* EGI (Eriksson, et al., 2004; Hui, et al., 2002). Although glycosylation of *Trichoderma reesei* cellulases is a complex issue,

our recombinant enzymes were found to be still active against both soluble and insoluble substrates such as 4-MUC, CMC and viscose.

Our recombinant enzymes, EGI and EGI\_L5, exhibited similar pH activity profiles and consistent temperature activity profiles to EGI produced by *Trichoderma reesei*. Cloning of EGI and EGI\_L5 in *Pichia pastoris* indicated that *Pichia pastoris* system can be used as a cloning host for endoglucanase 1. Both EGI and EGI\_L5 were expressed and produced actively in *Pichia pastoris*. Although different glycosylated forms were present, after affinity purification, both enzymes were obtained as single 70 kDa protein bands. Batch affinity purification with RAC is a promising affinity chromatography technique. RAC is an ultra high capacity adsorbent for cellulose binding domains. It has larger surface area thus a larger binding surface for the cellulose binding domains in comparison to Avicel. Because of this technique's relatively shorter process times (~30 minutes), cheaper price, and ease of use and reproducibility, it has come to be a frequently used research technique for the purification of cellulose binding domain tagged proteins (Hong, Wang, Ye, & Zhang, 2008). We successfully used the same approach to purify cellulases with cellulose binding domains in only one step. Fermentation products were directly applied on RAC without further processing. With the loop mutation, we obtained an active and stable endoglucanase which can be used for crosslinking and immobilization purposes. Loops that are inserted with this study had several Lysine residues which would create local high affinity points for crosslinking. Since the loop domain contains positively charged Lysine residues, it can be hypothesized that the mutant enzyme could be crosslinked from the  $\epsilon$ -amino groups of those residues effectively. Moreover, the loop domain was positioned away from the active site cleft, in order not to hinder the active site upon crosslinking.

### ***5.2. Effect of codon optimization on the production of EGI in Pichia pastoris***

*Trichoderma reesei* cellulase system has been known to exhibit its maximum activity around pH 5 (Bommarius, et al., 2008) and around 50 °C to 60 °C (our group's unpublished results). Our expression results of EGI and EGIs in *P. pastoris* were consistent with these data. Moreover, EGI catalytic domain expressed in *E. coli* has exhibited its maximum activity at pH 5 and has shown almost no activity at pH3 and pH7 (Nakazawa, et al., 2008). Heterologous expression of the codon optimized *eglI* gene and its native form in *P. pastoris* produced glycosylated enzyme products. Several



studies indicated heterogenous glycosylation of *Trichoderma reesei* EGI (Eriksson, et al., 2004; Hui, et al., 2002). Although glycosylation of *Trichoderma reesei* cellulases is a complex issue, our recombinant enzymes were found to be still active against substrates such as 4-MUC and CMC. Moreover, endoglucanases were produced as active and stable biocatalysts in *P. pastoris*. Penttila et al. expressed *Trichoderma reesei* endoglucanase 1 in *S. cerevisiae* (M. E. Penttila, et al., 1987) the recombinant protein was also found to be larger than the native protein produced by *T. reesei*. This was due to the differences between N-glycosylation patterns of *S. cerevisiae* and *T. reesei*.

Cloning of EGI and EGIs in *P. pastoris* indicated that *P. pastoris* system can be used as a suitable cloning host for endoglucanase 1. Both EGI and EGIs were expressed and produced actively in *P. pastoris*. As a result of codon optimization, recombinant endoglucanase 1 expression in *P. pastoris* is improved to 1.24 fold. It is known that increased GC-content prolongs mRNA half life and integrity in *P. pastoris* (Outchkourov, Stiekema, & Jongsma, 2002; Woo, et al., 2002). However it is also known that due to codon bias, there should be a careful balance between codon optimization and GC content optimization. In our case, *eglI* GC content was higher than the average GC content of *P. pastoris* and *eglI* gene was containing negative cis acting sites such as AT and GC-rich stretches etc. which may negatively influence expression. Our synthetic gene was carefully optimized to overcome these problems. GC content was reduced to obtain a similar value as of *P. pastoris* average GC content. Codon usage was biased to *P. pastoris* resulting in a high CAI value of 0.91. CAI index >0.9 is accepted as good for expression. Codon optimization of *eglI* gene improved endoglucanase 1 production in *P. pastoris*.

Codon optimization of the gene did not affect the glycosylation pattern as expected. Moreover, batch productions indicated an increase in EGI activity with time as the productivity based on the total protein concentration decreased. However the native enzyme EGI did not exhibit exactly the same trend. Total protein productivity of EGI clone has increased for the first 70 hours and then decreased while the enzyme activity has increased continuously. Our research indicated that *Pichia* system can also be used for expression and production of other cellulase components of *Trichoderma reesei*. Moreover, one-step affinity purification using regenerated amorphous cellulose can be easily applied to other cellulase components cloned in *P. pastoris*.

### 5.3. Cloning and Production of Recombinant Cellulases in *P. pastoris*

Bicatalytic EGI was prepared by overlap PCR extension in order to obtain an enzyme with increased activity and increased size. The recombinant enzyme produced in *P. pastoris* actively. But zymogram analysis revealed that the enzyme is degraded into catalytically active two components. Glycosylated forms of the enzyme were also present in zymogram gels. In order to alleviate the degradation problem by Pichia proteases fermentation conditions were changed by increasing the pH and decreasing the enzyme expression temperature. None of the conditions prevented the degradation of the enzyme. The degradation is thought to be arisen as a result of exposed linker region which connects the second catalytic domain to EGI domain. It is known that the linker region is o-glycosylated and this glycosylation reduces the affinity of proteases. If the linker region is not properly glycosylated, the enzyme would be open to the attack of Pichia proteases. Recombinant EGI\_BC has exhibited increased activity against 4-MUC and CMC since it has two catalytic domains.

EGIII and CBHI enzymes of *T. reesei* were expressed in *P. pastoris* as active enzyme products. Both enzymes have exhibited activity towards CMC. Activity of EGIII enzyme was also lower than EGI as expected since EGIII does not have a cellulose binding domain which helps the enzyme to adsorb to cellulosic substrates. EGIII has exhibited its maximum activity at 75 °C and at pH 5 which is much higher than the reported activity of the enzyme. Nakazawa et al. showed that EGIII stability was decreased to almost 0 % upon incubation at 60 and 70 °C for 15 minutes (Nakazawa, et al., 2008). EGIII was expressed in *E.coli*. This thermostability of EGIII expressed in *P. pastoris* may be due to glycosylation by Pichia glycosylation machinery since *E.coli* is not capable of glycosylation. It is known that glycosylation may confer enzymes extra stability (Kim, Kim, Raines, & Lee, 2004; Tang, et al., 2001).

CBHI activity is lower towards CMC since this enzyme is known to exhibit lower activity towards amorphous cellulose. Boer et al. expressed *cbh1* gene in *P. pastoris* GS115 strain under the control of alcohol oxidase (AOX1) and the glyceraldehyde-3-phosphate dehydrogenase (GAP) promoters and obtained over-glycosylated enzyme products and part of the enzyme produced under the control of AOX1 were not

correctly folded (Boer, Teeri, & Koivula, 2000). Moreover they have found out that over-glycosylation of the enzyme did not affect its thermal stability. Our results are consistent with their data. CBHI enzyme has exhibited activity towards CMC over a broad range of temperatures (at 65 °C has shown its maximum activity). Even at higher temperatures such as 85 and 95 °C, recombinant CBHI retained 75-80 % of its activity. It has shown optimum activity around pH 5 and pH 6. CBHI produced by *T. reesei* and CBHI produced by *P. pastoris* GS115 strain are known to show their maximum activity at pH 5 and their maximum stability at 60 °C (Boer, et al., 2000). Our enzyme has shown maximum activity at pH6 and 65 °C which are higher than the reported values.

#### **5.4. CLEA and Biopolishing**

Formation of pills on the surface of the fabric gives fabrics an aesthetic appearance. Application of cellulases for the removal of the pills on the surface is widely used in the industrial processes. The enzymatic process is very convenient for the cotton fabrics but results in the loss of tensile strength in viscose fabrics. Another problem arises from the fact that most of the commercial cellulases are unable to increase pilling values of the viscose as other fabrics.

Viscose knitted fabrics are more prone to pilling than any of the fabrics because of their structure and fiber properties. The outer shell of the viscose fiber consists of amorphous cellulose separated by smaller ordered crystalline regions. Since the amorphous regions are more, the viscose fibers are more prone to attack by cellulases. Degradation of amorphous regions would provide easy access of the enzymes to the ordered crystalline regions. Since the crystalline regions are mainly responsible for the tensile strength along the fiber axis, the tensile strength of the fabric drops upon cellulase action because of the degradation of this highly ordered crystalline regions. This study has shown that CLEA of commercial cellulases and use of recombinant mono component cellulases (EGI, EGI\_L5, EGI\_BC) can be used to alleviate the problem of lost tensile strength and can also be used to improve pilling notes of the viscose knitted and viscose woven fabrics. CLEA of commercial cellulases, EGI and EGI\_L5 were found to increase pilling values of the viscose fabrics by 20 %. Preliminary studies have shown that application of CLEA prepared from EGI and EGI\_L5 had not improved the pilling notes of the fabrics most probably due to loss of activity which might be caused by factors such as steric occlusion, hydrophobic interactions etc. More detailed

experiments and increased quantities of EGI and EGI\_L5 are needed to optimize CLEA production from both recombinant enzymes.

One drawback of CLEA application for biopolishing is that, the CLEA particles have different sizes and batch to batch variation was present. This variation in size is overcome by size fractionation of the cLEA particles and evaluating their effects separately. CLEA with different sizes were shown to improve enzyme pilling notes and tensile strength of the fabrics. CLEA prepared from EGI and loop mutant EGI\_L5 were found to be active against CMC and since they were produced in lower amounts could not be applied to fabric samples. However both EGI and EGI\_L5 have improved the pilling notes of the viscose knitted fabrics.

## CHAPTER 6

### 6. CONCLUSION

This study has shown that *P. pastoris* is an efficient host for production of recombinant endoglucanase 1, endoglucanase 1 mutants, endoglucanase 3 and cellobiohydrolase 1 of *Trichoderma reesei*. The recombinant enzymes are produced as active and stable biocatalysts since the *P. pastoris* system provides the suitable glycosylation patterns for stable and active production of *T. reesei* endoglucanases and cellobiohydrolases. *P. pastoris* endoglucanase 1, endoglucanase 3 and cellobiohydrolase 1 expression profiles are comparable to the ones reported in the literature (Boer, et al., 2000; Nakazawa, et al., 2008; M. E. Penttila, et al., 1987). The recombinant enzymes exhibited similar activities towards soluble substrates such as CMC and 4-MUC. In addition to endoglucanase 1 expression in *P. pastoris*, the effect of codon optimization on the EGI expression was studied. One step affinity purification has proven to be a suitable and rapid method for the purification of cellulases with cellulose binding domains. Moreover, the codon optimized endoglucanase 1 gene of *T. reesei* expressed in *P. pastoris* has improved the enzyme yield by 24 %. The change is not significantly different. This is expected since most of the codons existing in the native gene are also frequently used in *P. pastoris*. Overall, *P. pastoris* has proven to be an efficient host for the production of *T. reesei* cellulases.

Use of *in silico* molecular modelling methods with site directed mutagenesis for creating a hotspot for directed crosslinking of EGI away from the active site was successful in terms of producing an active mutant enzyme. Moreover, after crosslinking EGI\_L5 has shown similar activity as EGI. Their CLEA did not improve the pilling values but did not cause a reduction in fabric strength.

Preparation of crosslinked enzyme aggregates of a commercial cellulase and size fractionation of the CLEA particles alleviated the problem of pilling formation and tensile strength loss in viscose knitted and viscose woven fabrics. This effect was more pronounced in viscose woven fabrics. The pilling values were increased by 20 % upon

their application. Moreover, application of recombinant EGI, EGI\_L5 and EGI\_BC also improved the pilling notes of the viscose knitted fabrics.

## APPENDICES

### APPENDIX A

#### EQUIPMENTS

<b>Equipment</b>	<b>Brand Name/Model, Company</b>
Autoclave	Certoclav, Table Top Autoclave CV-EL-12L, AUSTRIA
	Hirayama, Hiclave HV-110, JAPAN
Balance	Sartorius, BP211D, GERMANY
	Sartorius, BP221S, GERMANY
	Sartorius, BP610, GERMANY
	Schimadzu, Libror EB-3200 HU, JAPAN
Burette	Borucam, TURKEY
Centrifuge	Eppendorf, 5415C, GERMANY
	Eppendorf, 5415D, GERMANY
	Eppendorf, 5415R, GERMANY
	Hitachi, Sorvall Discovery 100 SE, USA
	Hitachi, Sorvall RC5C Plus, USA
	Kendro Lab. Prod., Heraeus Multifuge 3L, GERMANY
Dialysis Membrane	Sigma, Cellusept
Distilled Water	Millipore, Elix-S, FRANCE
	Millipore, MilliQ Academic, FRANCE
Electrophoresis	Biorad Inc., USA
Eppendorf tubes(1.5-2ml)	Eppendorf
Falcon tubes(15-50ml)	TPP
Freezer	-70 °C, Kendro Lab. Prod., Heraeus Hfu486 Basic, GERMANY
	-20 °C, Bosch, TURKEY
Glasswares	Schott Duran, GERMANY

Hybridization Oven	Model 1012, Biolab, TURKEY
Ice Machine	Scotsman Inc., AF20, USA
Incubator	Memmert, Modell 300, GERMANY
	Memmert, Modell 600, GERMANY
Lyophilizer	
Magnetic Stirrer	ARE Heating Magnetic Stirrer, VELP Scientifica, ITALY
	Microstirrer, VELP Scientifica, ITALY
Micropipette	Eppendorf
Microscope	Olympos
Microtiter Plates (96-well)	TPP
Microtiterplate reader	Model 680, BioRad,
Microvave oven	Bosch, TURKEY
Multitube rotator	Labline
pH-meter	FisherBrand
Pipettoman	Hirschman Laborgate,
Power Supply	Biorad, PowerPac 300, USA
	Wealtec, Elite 300, USA
Refrigerator (+4°C)	Bosch, TURKEY
SDS-PAGE Gel Casting Apparatus	Biorad
SEM	Gemini 35 VP, Carl Zeiss, GERMANY
Shaker	Forma Scientific, Orbital Shaker 4520, USA
	C25HC Incubator shaker New Brunswick Scientific, USA
	GFL, Shaker 3011, USA
	New Brunswick Sci., Innova™ 4330, USA
Spectrophotometers	Schimadzu, UV-1208, JAPAN
	Schimadzu, UV-3150, JAPAN
	BioRad
Speed Vacuum	Savant, Speed Vac® Plus Sc100A, USA
	Savant, Refrigerated Vapor Trap RVT 400, USA
Tips	TPP
Thermal Heater	Bioblock Scientific
Thermomixer	Eppendorf
Water bath	Huber, Polystat cc1, GERMANY



## APPENDIX B

### Protein Sequences of Loop Models

```
>P1;EGI
structureX:1EG1: 2 :A:+370 :A:undefined:undefined:-1.00:-1.00
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCT
VNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSG
GYSSVSPRLYLSDSDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQM
DENGGA-----NQYNTAGANYGSGYCDAQCPVQTRNGTLNTSHQ
GFCCNEMDILEGNSRANALTPHSC TATAACDSAGCGFNPYSGYKSYGPG
DTVDTSKTFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI
SSCPSASAYGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSS
TEGNPSNILANNPNTHVVFSNIRWGDIGSTT*

>P1;EGI_L1
sequence:1EG1Loop1: 1 ::380 :endoglucanaseLoop1:Trichoderma
reesei mutant:::
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCT
VNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSG
GYSSVSPRLYLSDSDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQM
DENGGAKKGGKKKGGKNQYNTAGANYGSGYCDAQCPVQTRNGTLNTSHQ
GFCCNEMDILEGNSRANALTPHSC TATAACDSAGCGFNPYSGYKSYGPG
DTVDTSKTFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI
SSCPSASAYGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSS
TEGNPSNILANNPNTHVVFSNIRWGDIGSTT*

>P1;EGI_L2
sequence:1EG1L2: 1 ::380 :endoglucanaseL2:Trichoderma reesei
mutant:::
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCT
VNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSG
GYSSVSPRLYLSDSDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQM
DENGGAKKGGKKKGGKNQYNTAGANYGSGYCDAQCPVQTRNGTLNTSHQ
GFCCNEMDILEGNSRANALTPHSC TATAACDSAGCGFNPYSGYKSYGPG
DTVDTSKTFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI
SSCPSASAYGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSS
TEGNPSNILANNPNTHVVFSNIRWGDIGSTT*

>P1;EGI_L3
sequence:1EG1L3: 1 ::380 :endoglucanaseL3:Trichoderma reesei
mutant:::
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCT
VNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSG
GYSSVSPRLYLSDSDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQM
DENGGAKKGGKKKGGKNQYNTAGANYGSGYCDAQCPVQTRNGTLNTSHQ
GFCCNEMDILEGNSRANALTPHSC TATAACDSAGCGFNPYSGYKSYGPG
DTVDTSKTFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI
SSCPSASAYGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSS
TEGNPSNILANNPNTHVVFSNIRWGDIGSTT*

>P1;EGI_L4
sequence:1EG1L4: 1 ::380 :endoglucanaseL4:Trichoderma reesei
mutant:::
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCT
VNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSG
GYSSVSPRLYLSDSDGEYVMLKLNQELSFVDVLSALPCGENGSLYLSQM
DENGGAKKGGKKKGGKNQYNTAGANYGSGYCDAQCPVQTRNGTLNTSHQ
```

GFCCNEMDILEGNSRANALTPHSC TATA CDSAGCGFNPYGSYKSYYGPG  
DTVDTSKTF TIIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI  
SSCPSASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLDSGNAGPCSS  
TEGNPSNILANNPNTHVVF SNIRWGDIGSTT\*

>P1; EGI\_L5  
sequence:1EG1L5: 1 ::380 :endoglucanaseL5:Trichoderma reesei  
mutant:::  
QPGTSTPEVHPKLT TYKCTKSGGCVAQDTSVVL DWNYRWMHDANYNSCT  
VNGGVNTTLC PDEATCGKNC FIEGVDYAASGVTTSGSSLTMNQYMPSSSG  
GYSSVSPRLYLLD **KKGKKGKKGK**SDGEYVMLKLNQEL SFDVDLSALPCG  
ENGLYLSQMDENGGANQYNTAGANYGSGYCDAQCPVQ TWRNGTLN TSHQ  
GFCCNEMDILEGNSRANALTPHSC TATA CDSAGCGFNPYGSYKSYYGPG  
DTVDTSKTF TIIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI  
SSCPSASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLDSGNAGPCSS  
TEGNPSNILANNPNTHVVF SNIRWGDIGSTT\*

>P1; EGI\_L6  
sequence:1EG1L6: 1 ::380 :endoglucanaseL6:Trichoderma reesei  
mutant:::  
QPGTSTPEVHPKLT TYKCTKSGGCVAQDTSVVL DWNYRWMHDANYNSCT  
VNGGVNTTLC PDEATCGKNC FIEGVDYAASGVTTSGSSLTMNQYMPSSSG  
GYSSVSPRLYLLD **KKGKKGKKGK**SDGEYVMLKLNQEL SFDVDLSALPCG  
ENGLYLSQMDENGGANQYNTAGANYGSGYCDAQCPVQ TWRNGTLN TSHQ  
GFCCNEMDILEGNSRANALTPHSC TATA CDSAGCGFNPYGSYKSYYGPG  
DTVDTSKTF TIIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI  
SSCPSASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLDSGNAGPCSS  
TEGNPSNILANNPNTHVVF SNIRWGDIGSTT\*

>P1; EGI\_L7  
sequence:1EG1L7: 1 ::380 :endoglucanaseL7:Trichoderma reesei  
mutant:::  
QPGTSTPEVHPKLT TYKCTKSGGCVAQDTSVVL DWNYRWMHDANYNSCT  
VNGGVNTTLC PDEATCGKNC FIEGVDYAASGVTTSGSSLTMNQYMPSSSG  
GYSSVSPRLYLLD **KKGKKGKKGK**SDGEYVMLKLNQEL SFDVDLSALPCG  
ENGLYLSQMDENGGANQYNTAGANYGSGYCDAQCPVQ TWRNGTLN TSHQ  
GFCCNEMDILEGNSRANALTPHSC TATA CDSAGCGFNPYGSYKSYYGPG  
DTVDTSKTF TIIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTI  
SSCPSASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLDSGNAGPCSS  
TEGNPSNILANNPNTHVVF SNIRWGDIGSTT\*

## Protein Sequences of Cloned Genes

>EGI

MAPSVTLPLTTAILAIARLVAAQQPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNY  
RWMHDANYNSCTVNGGVNTTLCPEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSS  
SGGYSSVSPRLYLLDSDGEYVMLKLNQELSFVDVLSALPCGENGLYLSQMDENGGANQ  
YNTAGANYGSGYCDAQCPVQTRWRNGLNLSHQGFCCNEMDILEGNSRANALTPHSCTATA  
CDSAGCGFNPYGSYKSYGPGDVTDTSKFTIITQFNTDNGSPSGNLVSI TRKYQQNGV  
DIPSAQPGGDTISSCPASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPC  
SSTEGNPSNILANNPNTHVVF SNIRWGDIGSTTNSTAPPPPPASSTTFSTTRRSSTTSSS  
PSCTQTHWGQCGGIGYSGCKTCTSGTTCQYSNDYYSQCL

>CBHI

MYRKLAVISAFLATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVIDANWRWTHA  
TNSSTNCYDGNWSSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNSLSIGFVTQSAQKN  
VGARLYLMSADTTYQEFTLLGNEFSFDVDVSQLPCGLNGALYFVSMADGGVSKYPTNTA  
GAKYGTGYCDSQCPRDLKFINQANVEGWEPSSNNANTGIGGHGSCCSEMDIWEANSISE  
ALTPHPCTTVGQEICEGDGC GGTYSDNRYGGTCDPDGCDWNPYRLGNTSFYGPSSFTLD  
TTKCLTVVTQFETSGAINRYVQNGVTFQQPNAELGSYSGNELNDDYCTAEAEAFGGSSF  
SDKGLTQFKKATSGGMVLVMSLWDDYYANMLWLDSTYPTNETSSTPGAVRGCSTSSGV  
PAQVESQSPNAKVTF SNIKFGPIGSTGNPSSGNPPGGNRGTTTTTRRPATTTGSSPGPTQS  
HYGQCGGIGYSGPTVCASGTTCCQVLNPYYSQCL

>EGIII

MKFLQVLPALIPAALAQTS CDQWATFTGNGYTVSNNLWGASAGSGFGCVTAVSLSGGASW  
HADWQWSGGQNNVKS YQNSQIAIPQKRTVNSISSMPTTASWSYSGSNIRANVAYDLFTAA  
NPNHVITYSGDYELMIWLKYGDIGPIGSSQGT VNVGGQSWTLYYGYNGAMQVYSFVAQTN  
TTNYSGDVKNFFNYLRDNKGYNAAGQYVLSYQFGTEPFTGSGTLNVA SWTASIN

>EGI\_BC

QQPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCTVNGGVNTT  
LCPDEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSGGYSSVSPRLYLLDSDG  
EYVMLKLNQELSFVDVLSALPCGENGLYLSQMDENGGANQYNTAGANYGSGYCDAQ  
CPVQTRWRNGLNLSHQGFCCNEMDILEGNSRANALTPHSCTATA CDSAGCGFNPYGSY  
KSYGPGDVTDTSKFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTISS  
CPASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSSTEGNPSNI  
LANNPNTHVVF SNIRWGDIGSTTNSTAPPPPPASSTTFSTTRRSSTTSSSPCTQTHW  
GQCGGIGYSGCKTCTSGTTCQYSNDYYSQCLPPPPPPASSTTFSTTRRSSTTSSSPSCQ  
QPGTSTPEVHPKLTITYKCTKSGGCVAQDTSVVLWDWNYRWMHDANYNSCTVNGGVNTT  
LCPDEATCGKNCFIEGVDYAASGVTTSGSSLTMNQYMPSSSGGYSSVSPRLYLLDSDGE  
YVMLKLNQELSFVDVLSALPCGENGLYLSQMDENGGANQYNTAGANYGSGYCDAQ  
PVQTRWRNGLNLSHQGFCCNEMDILEGNSRANALTPHSCTATA CDSAGCGFNPYGSY  
KSYGPGDVTDTSKFTIITQFNTDNGSPSGNLVSI TRKYQQNGVDIPSAQPGGDTISS  
SCPSASAYGGLATMGKALSSGMVLVFSIWNDNSQYMNWLD SGNAGPCSSTEGNPSNI  
LANNPNTHVVF SNIRWGDIGSTTNSTA

## REFERENCES

- Adcock, S. A., & McCammon, J. A. (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. [doi: 10.1021/cr040426m]. *Chemical Reviews*, 106(5), 1589-1615.
- Aho, S. (1991). Structural and functional analysis of *Trichoderma reesei* endoglucanase I expressed in yeast *Saccharomyces cerevisiae*. *FEBS Lett*, 291(1), 45-49.
- Aho, S., Arffman, A., & Korhola, M. (1996). *Saccharomyces cerevisiae* mutants selected for increased production of *Trichoderma reesei* cellulases. *Applied Microbiology and Biotechnology*, 46(1), 36-45.
- Arnold, F. H. (1996). Directed evolution: creating biocatalysts for the future. *Chem. Eng. Sci*, 51(23), 5091- 5102.
- Azevedo, H., Bishop, D., & Cavaco-Paulo, A. (2000). Effects of agitation level on the adsorption, desorption, and activities on cotton fabrics of full length and core domains of EGV (*Humicola insolens*) and CenA (*Cellulomonas fimi*). [doi: 10.1016/S0141-0229(00)00205-2]. *Enzyme and Microbial Technology*, 27(3-5), 325-329.
- Bayer, E. A., Lamed, R., & Himmel, M. E. (2007). The potential of cellulases and cellulosomes for cellulosic waste management. [doi: DOI: 10.1016/j.copbio.2007.04.004]. *Current Opinion in Biotechnology*, 18(3), 237-245.
- Bayram Akcapinar, G. (2005). *COVALENT MODIFICATION OF ENZYMES FOR TEXTILE PROCESSES*. Sabancı University, Istanbul.
- Baysal, C., & Atilgan, A. R. (2001). Coordination topology and stability for the native and binding conformers of chymotrypsin inhibitor 2. *Proteins*, 45(1), 62-70.
- Beck, D. A., & Daggett, V. (2004). Methods for molecular dynamics simulations of protein folding/unfolding in solution. *Methods*, 34(1), 112-120.
- Bhat, M. K. (2000). Cellulases and related enzymes in biotechnology. *Biotechnology Advances*, 18, 355-383.
- Bhat, M. K., & Bhat, S. (1997). Cellulose degrading enzymes and their potential industrial applications. *Biotechnology Advances* 15, 583-620.
- Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J. C., et al. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*, 14(10), 1023-1031.
- Boer, H., Teeri, T. T., & Koivula, A. (2000). Characterization of *Trichoderma reesei* cellobiohydrolase Cel7A secreted from *Pichia pastoris* using two different promoters. *Biotechnology and Bioengineering*, 69(5), 486-494.
- Bommarius, A. S., Katona, A., Cheben, S. E., Patel, A. S., Ragauskas, A. J., Knudson, K., et al. (2008). Cellulase kinetics as a function of cellulose pretreatment. *Metab Eng*, 10(6), 370-381.
- Bornscheuer, U. T., & Pohl, M. (2001). Improved biocatalysts by directed evolution and rational protein design. *Curr Opin Chem Biol*, 5(2), 137-143.
- Bower, B., Clarkson, K., Collier, K., Kellis, J., Kelly, M., & Larenas, E. (1998).

- Brown, R. M. (Ed.). (1982). *Cellulose and Other Natural Polymer Systems: Biogenesis, Structure and Degradation*, . New York and London: Plenum Press
- Busto, M. D., Ortega, N., & Perez-Mateos, M. (1997). Stabilisation of cellulases by cross-linking with glutaraldehyde and soil humates. [doi: DOI: 10.1016/S0960-8524(97)00001-1]. *Bioresource Technology*, 60(1), 27-33.
- Carter, P. (1986). Site-directed mutagenesis. *Biochem J*, 237(1), 1-7.
- Cavaco-Paulo, A., Almeida, L., & Bishop, D. (1996). Cellulase activities and finishing effects. *Text. Chem. Col.*, 28(6), 28-32.
- Chernoglazov, V. M., Jafarova, A. N., & Klyosov, A. A. (1989). Continuous photometric determination of endo-1,4-beta-D-glucanase (cellulase) activity using 4-methylumbelliferyl-beta-D-cellobioside as a substrate. *Anal Biochem*, 179(1), 186-189.
- Ciechańska, D., Struszczyk, H., Miettinen-Oinonen, A., & Strobin, G. (2002). Enzymatic Treatment of Viscose Fibres Based Woven Fabric. *FIBRES & TEXTILES in Eastern Europe* 6, 60-63.
- Daggett, V. (2006). Protein folding-simulation. *Chem Rev*, 106(5), 1898-1916.
- Day, R., Bennion, B. J., Ham, S., & Daggett, V. (2002). Increasing temperature accelerates protein unfolding without changing the pathway of unfolding. *J Mol Biol*, 322(1), 189-203.
- De Mori, G. M., Meli, M., Monticelli, L., & Colombo, G. (2005). Folding and misfolding of peptides and proteins: insights from molecular simulations. *Mini Rev Med Chem*, 5(4), 353-359.
- Divne, C., Stahlberg, J., Reinikainen, T., Ruohonen, L., Pettersson, G., Knowles, J., et al. (1994). The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei*. *Science*, 265(5171), 524-528.
- Eriksson, T., Stals, I., Collen, A., Tjerneld, F., Claeysens, M., Stalbrand, H., et al. (2004). Heterogeneity of homologously expressed *Hypocrea jecorina* (*Trichoderma reesei*) Cel7B catalytic module. *Eur J Biochem*, 271(7), 1266-1276.
- Fersht, A. R., & Daggett, V. (2002). Protein folding and unfolding at atomic resolution. *Cell*, 108(4), 573-582.
- Floriano, W. B., Domont, G. B., & Nascimento, M. A. (2007). A molecular dynamics study of the correlations between solvent-accessible surface, molecular volume, and folding state. *J Phys Chem B*, 111(7), 1893-1899.
- Ghose, T. K. (1987). Measurement of Cellulase Activities. *Pure & Applied Chemistry*, 59(2), 257-268.
- Glykos, N. M. (2006). Software news and updates. Carma: a molecular dynamics analysis program. *J Comput Chem*, 27(14), 1765-1768.
- Gruno, M., Valjamae, P., Pettersson, G., & Johansson, G. (2004). Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnol Bioeng*, 86(5), 503-511.
- Gu, W., Wang, T., Zhu, J., Shi, Y., & Liu, H. (2003). Molecular dynamics simulation of the unfolding of the human prion protein domain under low pH and high temperature conditions. *Biophys Chem*, 104(1), 79-94.
- Hakansson, U., Fagerstam, L., Pettersson, G., & Andersson, L. (1978). Purification and characterization of a low molecular weight 1,4-beta-glucan glucanohydrolase from the cellulolytic fungus *Trichoderma viride* QM 9414. *Biochim Biophys Acta*, 524(2), 385-392.

- Halliwell, G., & Vincent, R. (1981). The action on cellulose and its derivatives of a purified 1,4-beta-glucanase from *Trichoderma koningii*. *Biochem J*, 199(2), 409-417.
- Heikinheimo, L. (2002). *Trichoderma reesei* cellulases in processing of cotton
- Heikinheimo, L., & Buchert, J. (2001). Synergistic effects of *Trichoderma reesei* cellulases on the properties of knitted cotton fabric. *Textile Research Journal*, 71, 672-677.
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, 280 ( Pt 2), 309-316.
- Henrissat, B., & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*, 293 ( Pt 3), 781-788.
- Hong, J., Wang, Y., Ye, X., & Zhang, Y. H. (2008). Simple protein purification through affinity adsorption on regenerated amorphous cellulose followed by intein self-cleavage. *J Chromatogr A*, 1194(2), 150-154.
- Hong, J., Ye, X., Wang, Y., & Zhang, Y. H. (2008). Bioseparation of recombinant cellulose-binding module-proteins by affinity adsorption on an ultra-high-capacity cellulosic adsorbent. *Anal Chim Acta*, 621(2), 193-199.
- Huang, X., Gao, D., & Zhan, C.-G. (2011). Computational design of a thermostable mutant of cocaine esterase via molecular dynamics simulations. *Organic & Biomolecular Chemistry*, 9(11), 4138-4143.
- Hui, J. P. M., White, T. C., & Thibault, P. (2002). Identification of glycan structures and glycosylation sites in cellobiohydrolase II and endoglucanase I and II from *Trichoderma reesei*. *Glycobiology* 12, 837-849.
- Hung, H. C., Chen, Y. H., Liu, G. Y., Lee, H. J., & Chang, G. G. (2003). Equilibrium protein folding-unfolding process involving multiple intermediates. *Bull Math Biol*, 65(4), 553-570.
- Karplus, M., & Sali, A. (1995). Theoretical studies of protein folding and unfolding. *Curr Opin Struct Biol*, 5(1), 58-73.
- Kiernan, J. (2000). Formaldehyde, formalin, paraformaldehyde and glutaraldehyde: What they are and what they do. *Microscopy Today*, 00-1, 8-12.
- Kim, B.-M., Kim, H., Raines, R. T., & Lee, Y. (2004). Glycosylation of onconase increases its conformational stability and toxicity for cancer cells. [doi: 10.1016/j.bbrc.2004.01.153]. *Biochemical and Biophysical Research Communications*, 315(4), 976-983.
- Klepeis, J. L., Lindorff-Larsen, K., Dror, R. O., & Shaw, D. E. (2009). Long-timescale molecular dynamics simulations of protein structure and function. [doi: 10.1016/j.sbi.2009.03.004]. *Current Opinion in Structural Biology*, 19(2), 120-127.
- Kleywegt, G. J., Zou, J.-Y., Divne, C., Davies, G. J., Sinning, I., Ståhlberg, J., et al. (1997). The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes. [doi: 10.1006/jmbi.1997.1243]. *Journal of Molecular Biology*, 272(3), 383-397.
- Kumar, A., Yoon, M., & Purtell, C. (1997). Optimizing the Use of Cellulase Enzymes in Finishing Cellulosic Fabrics *Textile Chemist and Colorist* 29, 37-42.
- Lamed, R., & Bayer, E. A. (1988). The Cellulosome of *Clostridium thermocellum*. In L. L. Allen (Ed.), *Advances in Applied Microbiology* (Vol. Volume 33, pp. 1-46): Academic Press.

- Lameira, J. n., Alves, C. u. N., Tuñón, I. a., Martí, S., & Moliner, V. (2011). Enzyme Molecular Mechanism as a Starting Point to Design New Inhibitors: A Theoretical Study of O-GlcNAcase. [doi: 10.1021/jp202079e]. *The Journal of Physical Chemistry B*, 115(20), 6764-6775.
- Lenting, H. B. M., & Warmoeskerken, M. M. C. G. (2001). Guidelines to come to minimized tensile strength loss upon cellulase application. *Journal of Biotechnology* 89, 227-232.
- Liming, X., & Xueliang, S. (2004). High-yield cellulase production by *Trichoderma reesei* ZU-02 on corn cob residue. *Bioresour Technol*, 91(3), 259-262.
- Liu, H. L., & Wang, W. C. (2003a). Molecular dynamics simulations of the unfolding of the starch binding domain from *Aspergillus niger* glucoamylase. *J Biomol Struct Dyn*, 20(5), 615-622.
- Liu, H. L., & Wang, W. C. (2003b). Predicted unfolding order of the 13 alpha-helices in the catalytic domain of glucoamylase from *Aspergillus awamori* var. X100 by molecular dynamics simulations. *Biotechnol Prog*, 19(5), 1583-1590.
- Liu, H. L., Wang, W. C., & Hsu, C. M. (2003). Molecular dynamics simulations of the unfolding mechanism of the catalytic domain from *Aspergillus awamori* var. X100 glucoamylase. *J Biomol Struct Dyn*, 20(4), 567-574.
- Liu, J., Otto, E., Lange, N., Husain, P., Condon, B., & Lund, H. (2000). Selecting cellulases for Bio-polishing based on enzyme selectivity and process conditions. *. Textile Chemist and Colorist & American Dyestuff Reporter*, 32(5), 30-36.
- López-Serrano, P., Cao, L., van Rantwijk, F., & Sheldon, R. A. (2002). Cross-linked enzyme aggregates with enhanced activity: application to lipases. *Biotechnology Letters*, 24(16), 1379-1383.
- Mark, A. E., & van Gunsteren, W. F. (1992). Simulation of the thermal denaturation of hen egg white lysozyme: trapping the molten globule state. *Biochemistry*, 31(34), 7745-7748.
- Medve, J., Karlsson, J., Lee, D., & Tjerneld, F. (1998). Hydrolysis of microcrystalline cellulose by cellobiohydrolase I and endoglucanase II from *Trichoderma reesei*: adsorption, sugar production pattern, and synergism of the enzymes. *Biotechnol Bioeng*, 59(5), 621-634.
- Miettinen-Oinonen, A., Paloheimo, M., Lantto, R., & Suominen, P. (2005). Enhanced production of cellobiohydrolases in *Trichoderma reesei* and evaluation of the new preparations in biofinishing of cotton. *J Biotechnol*, 116(3), 305-317.
- Nakazawa, H., Okada, K., Kobayashi, R., Kubota, T., Onodera, T., Ochiai, N., et al. (2008). Characterization of the catalytic domains of *Trichoderma reesei* endoglucanase I, II, and III, expressed in *Escherichia coli*. *Appl Microbiol Biotechnol*, 81(4), 681-689.
- Nakazawa, H., Okada, K., Onodera, T., Ogasawara, W., Okada, H., & Morikawa, Y. (2009). Directed evolution of endoglucanase III (Cel12A) from *Trichoderma reesei*. *Appl Microbiol Biotechnol*, 83(4), 649-657.
- Németh, A., Kamondi, S., Szilágyi, A., Magyar, C., Kovári, Z., & Závodszy, P. (2002). Increasing the thermal stability of cellulase C using rules learned from thermophilic proteins: a pilot study. [doi: 10.1016/S0301-4622(02)00027-3]. *Biophysical Chemistry*, 96(2-3), 229-241.
- Okada, H., Mori, K., Tada, K., Nogawa, M., & Morikawa, Y. (2000). Identification of active site carboxylic residues in *Trichoderma reesei* endoglucanase Cel12A by site-directed mutagenesis. [doi: 10.1016/S1381-1177(00)00137-5]. *Journal of Molecular Catalysis B: Enzymatic*, 10(1-3), 249-255.

- Outchkourov, N. S., Stiekema, W. J., & Jongsma, M. A. (2002). Optimization of the expression of equistatin in *Pichia pastoris*. *Protein Expr Purif*, 24(1), 18-24.
- Penttila, M., Nevalainen, H., Ratto, M., Salminen, E., & Knowles, J. (1987). A versatile transformation system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene*, 61(2), 155-164.
- Penttila, M. E., Andre, L., Saloheimo, M., Lehtovaara, P., & Knowles, J. K. (1987). Expression of two *Trichoderma reesei* endoglucanases in the yeast *Saccharomyces cerevisiae*. *Yeast*, 3(3), 175-185.
- Percival Zhang, Y. H., Himmel, M. E., & Mielenz, J. R. (2006). Outlook for cellulase improvement: Screening and selection strategies. [doi: DOI: 10.1016/j.biotechadv.2006.03.003]. *Biotechnology Advances*, 24(5), 452-481.
- Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., et al. (2005). Scalable molecular dynamics with NAMD. *J Comput Chem*, 26(16), 1781-1802.
- Qin, Y., Wei, X., Liu, X., Wang, T., & Qu, Y. (2008). Purification and characterization of recombinant endoglucanase of *Trichoderma reesei* expressed in *Saccharomyces cerevisiae* with higher glycosylation and stability. *Protein Expr Purif*, 58(1), 162-167.
- Reinikainen, T., Teeman, O., & Teeri, T. T. (1995). Effects of pH and high ionic strength on the adsorption and activity of native and mutated cellobiohydrolase I from *Trichoderma reesei*. *Proteins*, 22(4), 392-403.
- Richards, F. M., & Knowles, J. R. (1968). Glutaraldehyde as a protein cross-linkage reagent. *J Mol Biol*, 37(1), 231-233.
- Rignall, T. R., Baker, J. O., McCarter, S. L., Adney, W. S., Vinzant, T. B., Decker, S. R., et al. (2002). Effect of single active-site cleft mutation on product specificity in a thermostable bacterial cellulase. *Appl Biochem Biotechnol*, 98-100, 383-394.
- Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*, 234(3), 779-815.
- Saloheimo, M., Nakari-Setälä, T., Tenkanen, M., & Penttila, M. (1997). cDNA cloning of a *Trichoderma reesei* cellulase and demonstration of endoglucanase activity by expression in yeast. *Eur J Biochem*, 249(2), 584-591.
- Sandgren, M., Stahlberg, J., & Mitchinson, C. (2005). Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. *Prog Biophys Mol Biol*, 89(3), 246-291.
- Schmidt, M., Bottcher, D., & Bornscheuer, U. T. (2009). Protein engineering of carboxyl esterases by rational design and directed evolution. *Protein Pept Lett*, 16(10), 1162-1171.
- Schoevaart, R., Wolbers, M. W., Golubovic, M., Ottens, M., Kieboom, A. P., van Rantwijk, F., et al. (2004). Preparation, optimization, and structures of cross-linked enzyme aggregates (CLEAs). *Biotechnol Bioeng*, 87(6), 754-762.
- Shao, Z., & Arnold, F. H. (1996). Engineering new functions and altering existing functions. *Curr Opin Struct Biol*, 6(4), 513-518.
- Sheldon, R. A., Schoevaart, R., & Langen, L. M. (2006). Cross-Linked Enzyme Aggregates. In J. M. Guisan (Ed.), *Immobilization of Enzymes and Cells* (Vol. 22, pp. 31-45): Humana Press.
- Srisodsuk, M., Reinikainen, T., Penttila, M., & Teeri, T. T. (1993). Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose. *J Biol Chem*, 268(28), 20756-20761.



- Ståhlberg, J. (1991). *Functional organization of cellulases from Trichoderma reesei*. Unpublished Doctoral thesis, Uppsala University, Uppsala.
- Tang, S.-J., Shaw, J.-F., Sun, K.-H., Sun, G.-H., Chang, T.-Y., Lin, C.-K., et al. (2001). Recombinant Expression and Characterization of the *Candida rugosa* lip4 Lipase in *Pichia pastoris*: Comparison of Glycosylation, Activity, and Stability. [doi: 10.1006/abbi.2000.2235]. *Archives of Biochemistry and Biophysics*, 387(1), 93-98.
- Ulker, A., & Sprey, B. (1990). Characterization of an unglycosylated low molecular weight 1,4-beta-glucan-glucanohydrolase of *Trichoderma reesei*. *FEMS Microbiol Lett*, 57(3), 215-219.
- Vallejo, A. N., Pogulis, L. J., & Pease, L. R. (1994). In vitro synthesis of novel genes: mutagenesis and recombination by PCR. *Genome Research*, 4, 123-130.
- van Gunsteren, W. F., & Mark, A. E. (1992a). On the interpretation of biochemical data by molecular dynamics computer simulation. *Eur J Biochem*, 204(3), 947-961.
- van Gunsteren, W. F., & Mark, A. E. (1992b). Prediction of the activity and stability effects of site-directed mutagenesis on a protein core. *J Mol Biol*, 227(2), 389-395.
- Videb, A. B. T. K. H., DK), Andersen, Lars Dalg Ang Rd (Virum, DK). (2000). United States Patent No.
- Videbaek, T., & Andersen, L. D. (1993).
- Vinzant, T. B., Adney, W. S., Decker, S. R., Baker, J. O., Kinter, M. T., Sherman, N. E., et al. (2001). Fingerprinting *Trichoderma reesei* hydrolases in a commercial cellulase preparation. *Appl Biochem Biotechnol*, 91-93, 99-107.
- Wilson, D. B. (2009). Cellulases and biofuels. [doi: DOI: 10.1016/j.copbio.2009.05.007]. *Current Opinion in Biotechnology*, 20(3), 295-299.
- Woo, J. H., Liu, Y. Y., Mathias, A., Stavrou, S., Wang, Z., Thompson, J., et al. (2002). Gene optimization is necessary to express a bivalent anti-human anti-T cell immunotoxin in *Pichia pastoris*. *Protein Expr Purif*, 25(2), 270-282.
- Wu, S., & Letchworth, G. J. (2004). High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *Biotechniques*, 36(1), 152-154.
- Yao, Q., Sun, T. T., Liu, W. F., & Chen, G. J. (2008). Gene cloning and heterologous expression of a novel endoglucanase, swollenin, from *Trichoderma pseudokoningii* S38. *Biosci Biotechnol Biochem*, 72(11), 2799-2805.
- You, L., & Arnold, F. H. (1996). Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. *Protein Eng*, 9(1), 77-83.
- Yuan, X.-y., Shen, N.-x., Sheng, J., & Wei, X. (1999). Immobilization of cellulase using acrylamide grafted acrylonitrile copolymer membranes. [doi: DOI: 10.1016/S0376-7388(98)00313-5]. *Journal of Membrane Science*, 155(1), 101-106.
- Zhang, Y. H., & Lynd, L. R. (2004). Toward an Aggregated Understanding of Enzymatic Hydrolysis of Cellulose: Noncomplexed Cellulase Systems, Review. *Biotechnology and Bioengineering* 88(7), 797-824.