

CLONING, EXPRESSION, AND IN SILICO ANALYSES OF SELECTED XTH
ENZYMES FROM DIFFERENT PLANT SPECIES



by
Selin Mumcu

Submitted to Graduate School of Natural and Applied Sciences
in Partial Fulfillment of the Requirements
for the Degree of Master of Science in
Biotechnology

Yeditepe University
2022

CLONING, EXPRESSION, AND IN SILICO ANALYSES OF SELECTED XTH
ENZYMES FROM DIFFERENT PLANT SPECIES

APPROVED BY:

Assist. Prof. Dr. Andrew John Harvey
(Thesis Supervisor)
(Yeditepe University)

Assist. Prof. Dr. Bahar Soğutalmaz Özdemir
(Yeditepe University)

Assist. Prof. Dr. Merve Seven
(Bahçeşehir University)

DATE OF APPROVAL:/....../2022

I hereby declare that this thesis is my own work and that all information in this thesis has been obtained and presented in accordance with academic rules and ethical conduct. I have fully cited and referenced all material and results as required by these rules and conduct, and this thesis study does not contain any plagiarism. If any material used in the thesis requires copyright, the necessary permissions have been obtained. No material from this thesis has been used for the award of another degree.

I accept all kinds of legal liability that may arise in case contrary to these situations.

Name, Last name

Selin Mumcu

Signature

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor, Assist. Prof. Dr. Andrew J. Harvey for guiding and assisting me with his experience and knowledge in the field in which I'm most willing to work. I am sincerely grateful for him to always be ready for keeping a supportive and motivating attitude throughout my thesis.

During my master's degree, I had the opportunity to work with amazing members of Plant Lab and Bioprocess Lab. I would like to thank my senior, Dr. Mehtap Aydın for being my lifesaver at the laboratory; my jury members for their academic knowledge, moral support, and helpful remarks Assist. Prof. Dr. Bahar Soğutalmaz Özdemir and Assist. Prof. Dr. Merve Seven. I want to thank my friends, Deniz Yeşil Baysal and Naz Ünsal who make me feel like I'm at home even though I'm in the middle of a chaotic experiment; my ride or die, Gizem Özer; and despite miles of the distance she supported me and patiently waited for me to come together, Elif Gugu Karaca; Moreover, I would like to thank my “before anyone else”, Mehmet Can Sucu for all support and vision that changed my life.

Last but not least, I am immensely thankful to my beloved mother for making me who I am today and my family, mostly to my grandmother for being my housemate and my biggest supporter.

ABSTRACT

CLONING, EXPRESSION, AND IN SILICO ANALYSES OF SELECTED XTH ENZYMES FROM DIFFERENT PLANT SPECIES

During the plant life cycle, the plant cell wall undergoes alterations and adaptations which are accomplished by enzymes able to modify cell wall polysaccharides. A significant and abundant non-cellulosic β -linked polysaccharide called xyloglucan binds cellulose microfibrils to create supporting structures in the cell wall of the plant. Modifications of xyloglucan in the plant cell wall are carried out by the crucial modifying enzyme xyloglucan endotransglycosylase/hydrolases (XTHs) which are divided into six phylogenetic groups (Group I, II, III-A, III-B, IV, and EG16) depending on the similarities of their amino acid sequences. Understanding the roles of these enzymes that change the carbohydrate matrix, is essential for altering cell wall characteristics to increase yield and preserve the plant from external factors. To shed light on the working process of XTH in different species and elucidate the relationships between the plants and their polysaccharides substrates, we selected the *AtXTH30*, *Arabidopsis thaliana* XTH gene which belongs to the poorly studied sub-clade Group III-B; and *EtXTH1*, *Equisetum telmateia* XTH gene, which has not been enzymatically characterized. The objective of this thesis was to investigate *A. thaliana* and *Equisetum* XTH enzymes at the molecular, transcriptomic, and protein levels to gain a better insight into how these enzymes work and what functions they serve. Unfortunately, *EtXTH1* and *AtXTH30* enzymes could not be heterologously expressed during the period of this study. However, XTH genes of *Arabidopsis thaliana*, and *Equisetum* family members *Equisetum telmateia*, *Equisetum hymale*, and *Equisetum diffusum* were identified by transcriptome analyses. Phylogenetic and transcriptome analyses revealed that catalytic motifs of XTH genes were maintained and, in certain circumstances, varied from species to species by amino acid substitutions, influencing the enzymatic activity of specific XTH genes. Highly conserved protein motifs were determined and the *AtXTH30* protein structure was modeled. Moreover, the generated phylogenetic tree provides information about the relationship between the XTH enzymes that could be broadly capable of transferase/hydrolase activity.

ÖZET

FARKLI BİTKİ TÜRLERİNDEN SEÇİLMİŞ XTH ENZİMLERİNİN KLONLAMA, EKSPRESYON VE İN SİLİKO ANALİZLERİ

Bitki hücre duvarı, bitkinin yaşam döngüsü boyunca hücre duvarı polisakkaritlerini modifiye edebilen enzimler tarafından gerçekleştirilen değişikliklere ve adaptasyonlara uğramaktadır. Ksiloglukan adındaki önemli ve bol miktarda bulunan selülozik olmayan β -bağlı bir polisakkarit, bitki hücre duvarında yük taşıyan yapılar oluşturmak için selüloz mikrofibrillerini bağlar. Bitki hücre duvarındaki modifikasyonlar, önemli bir modifiye edici enzim olan ksiloglukan endotransglikosilaz/hidrolazlar (XTH'ler) tarafından gerçekleştirilmektedir ve XTH genleri amino asit sekanslarındaki benzerliklere göre (Grup I, II, III-A, III-B, IV ve EG16) altı filogenetik gruba ayrılmaktadır. Karbonhidrat matrisini değiştiren bu enzimlerin rollerini anlamak; hücre duvarı özelliklerini modifiye ederek verimi artırmak ve bitkiyi dış etkenlerden korumak için gereklidir. XTH'nin farklı bitki türlerindeki çalışma mekanizmalarını aydınlatmak, ve bitkiler ile polisakkarit substratları arasındaki ilişkileri aydınlatmak için, *Arabidopsis thaliana* XTH geni olan ve yeterince aydınlatılmamış bir alt grup olan Grup III-B'ye ait *AtXTH30* genini, ve enzimatik karakterizasyonu henüz yapılmamış *Equisetum telmateia* XTH geni olan *EtXTH1* genini seçtik. Bu tezin amacı ise, *A. thaliana* ve *Equisetum* XTH enzimlerini moleküler, transkriptomik ve protein seviyelerinde araştırmak ve bu enzimlerin nasıl çalıştığını ve hangi işlevleri yerine getirdiğini daha iyi anlamaktır. *EtXTH1* ve *AtXTH30* enzimleri, bu çalışma süresince heterolog ekspresyon ile üretilenmemiştir. Ancak, *Arabidopsis thaliana*, ve *Equisetum* ailesi üyelerinden *E. telmateia*, *E. hymale* ve *E. diffusum* türlerine ait XTH genleri transkriptom analizleri ile tanımlanmıştır. Filogenetik ve transkriptom analizleri, XTH genlerinin katalitik motiflerinin korunduğunu ve belirli durumlarda, amino asit ikameleri ile türden türe değişiklik göstererek spesifik XTH genlerinin enzimatik aktivitesini etkilediğini ortaya çıkarmıştır. Çalışma sonunda yüksek oranda korunmuş protein motifleri belirlenmiştir ve *AtXTH30* protein yapısı modellenmiştir. Bu bilgilerin ışığında oluşturulan filogenetik ağaç, geniş ölçüde transferaz/hidrolaz aktivitesine sahip olabilen XTH enzimleri arasındaki ilişki hakkında bilgi sağlayabilmektedir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iv
ABSTRACT	v
ÖZET	vi
LIST OF FIGURES	x
LIST OF TABLES	xii
LIST OF SYMBOLS/ABBREVIATIONS	xiii
1. INTRODUCTION	1
1.1. ARABIDOPSIS THALIANA	1
1.2. Equisetum telmateia	2
1.3. PLANT CELL WALL	4
1.4. PLANT CELL WALL STRUCTURE	5
1.5. CELLULOSE	6
1.6. XYLOGLUCAN	8
1.7. MIXED LINKED BETA-GLUCANS	11
1.8. XYLAN	12
1.9. MANNAN	12
1.10. PECTIN	13
1.11. XYLOGLUCAN ENDOTRANSGLYCOSYLASE/HYDROLASE	13
1.12. POLYSACCHARIDE AND OLIGOSACCHARIDE SUBSTRATES	15
1.13. HETEROLOGOUS EXPRESSION AND PICHIA PASTORIS SYSTEM	15
1.14. XTH ENZYME CLASSIFICATION AND EVOLUTIONARY LINKAGES ...	19
1.15. AIM OF THE STUDY	20
2. MATERIALS	21
2.1. EQUIPMENT AND MATERIALS	21
2.2. CHEMICALS & DISPOSABLES	22
2.3. BIOLOGICAL MATERIAL	24
2.3.1. Reverse and Forward Primers	25

2.4. DONOR AND ACCEPTOR SUBSTRATES	22
3. METHODS	28
3.1. TARGET ENZYME SELECTION	41
3.2. GLYCEROL STOCK PREPARATION	28
3.3. EXPRESSION PLASMID PREPARATION	28
3.3 PLASMID ISOLATION	30
3.4. BACTERIAL TRANSFORMATION	30
3.5. PHENOL EXTRACTION AND ETHANOL PRECIPITATION OF DNA	30
3.6. COMPETENT PICHIA PASTORIS PREPARATION	31
3.8. PICHIA PASTORIS TRANSFORMATION AND SELECTION FOR POSITIVE COLONIES	31
3.9. COLONY PCR	32
3.10. SDS-PAGE AND WESTERN BLOTTING	33
3.11. HETEROLOGOUS PROTEIN PRODUCTION	34
3.12. PREPARATION OF DONOR AND ACCEPTOR MOLECULES	34
3.13. ENZYME CHARACTERIZATION AND ACTIVITY ANALYSIS	34
3.14. HPLC ANALYSIS	35
3.15. IN SILICO SEQUENCE ANALYSIS	35
3.15.1. Multiple Sequence Alignment and Phylogenetic Tree	38
3.15.2. Signal Peptide & N-linked Glycosylation Site Prediction & DREME	38
3.15.3. Swiss Modeling & Swiss PDB Viewer	38
4. RESULTS	41
4.1. TARGET ENZYME SELECTION	41
4.2. PROTEIN PRODUCTION AND CHARACTERIZATION	42
4.2.1. Expression Plasmid Preparation	42
4.2.2. Yeast Transformation	44
4.2.3. Positive Colony Selection and Screening	46
4.2.4. Detection of AtXTH30 Protein	47
4.3. IN SILICO ANALYSIS OF XTH GENES	49
4.3.1. Multiple Sequence Alignment (MSA)	49
4.3.2. Phylogenetic Tree of XTH Genes	56

4.3.3.	Signal Peptide and Mature Protein Determination	59
4.3.4.	DREME	60
4.3.5.	Protein Structure Homology Modeling	62
5.	DISCUSSION	65
6.	CONCLUSION	70
	REFERENCES	72



LIST OF FIGURES

Figure 1.1. Parts of the flowering plant <i>Arabidopsis thaliana</i>	1
Figure 1.2. Great horsetail, <i>Equisetum telmateia</i>	15
Figure 1.3. Phylogenetic tree of ancient and present <i>Equisetales</i>	16
Figure 1.4. Cell wall polysaccharide biosynthesis	19
Figure 1.5. Primary cell wall structure	20
Figure 1.6. Tamarind xyloglucan, XG7, XG8 and XG9 schematic representations.	22
Figure 1.7. A broad range of xyloglucan (XyG) sidechains.	23
Figure 1.8. Plasmid map of target DNA with expression vector pPICZ α -C	30
Figure 1.9. Evolution of EG16 licheninase-like ancestor XTH gene products	31
Figure 2.1. Representation of Tamarind Seed Xyloglucan (TXG) polysaccharide as donor substrate for enzymatic activity assays	39
Figure 4.1. Uncut plasmid DNA's of <i>EtXTH1</i> and <i>AtXTH30</i>	50
Figure 4.2. PCR products of <i>AtXTH30</i> on 1 percent agarose gel electrophoresis.....	51
Figure 4.3. The expression cassette of <i>Pichia pastoris</i> genome	52
Figure 4.4. <i>AtXTH30</i> -pPICZ α -C <i>E.coli</i> transformant selection	53
Figure 4.5. Sequence of the expression vector pPICZ α -C	53
Figure 4.6. Growth curve of <i>P. pastoris</i> cells containing <i>AtXTH30</i> -pPICZ α -C recombinant plasmids during methanol induction	55
Figure 4.7. SDS-PAGE image of <i>AtXTH30</i> proteins	56
Figure 4.8. Multiple Sequence Alignment of <i>A. thaliana</i> XTH Sequences	58

Figure 4.9. The Multiple Sequence Alignment of Equisetum XTH Sequences.....	62
Figure 4.10. Phylogenetic tree of XTH genes of <i>A. thaliana</i> , <i>E. telmateia</i> , <i>E. diffusum</i> , <i>E. hymale</i> , <i>E. fluviatile</i> and <i>O. sativa</i>	61
Figure 4.11. <i>AtXTH30</i> levels during flowering	62
Figure 4.12. Deduced amino acid sequence of <i>AtXTH30</i> contain signal peptide and mature proteins	62
Figure 4.13. Predicted Signal Peptide and Cleavage Site of <i>AtXTH30</i>	64
Figure 4.14. Predicted protein structure of <i>AtXTH30</i> by SWISS-MODEL	64
Figure 4.15. Predicted protein structure of <i>AtXTH30</i> with labeled conserved residues.....	60
Figure 4.16. Predicted protein structure of <i>AtXTH30</i> with green-colored catalytic site	64

LIST OF TABLES

Table 2.1. List of equipment and company names	22
Table 2.2. List of chemicals and disposables with company names supplied by.....	23
Table 2.3. List of biological materials with company names supplied by	25
Table 2.4. . Forward and reverse primer sequences for <i>EtXTH1</i>	26
Table 2.5. Specifications of forward <i>EtXTH1</i> primer	26
Table 2.6. Specifications of reverse <i>EtXTH1</i> primer.....	26
Table 2.7. List of donor and acceptor substrate, abbreviations, and catalog number	28
Table 2.8. List of acceptor substrate, abbreviation and catalog number	28
Table 3.1. Accession numbers of selected <i>A. thaliana</i> XTH genes	37
Table 4.1. DREME Results for <i>Equisetum</i> XTHs.....	61
Table 4.2. DREME Results for <i>Arabidopsis thaliana</i> XTHs	61

LIST OF SYMBOLS/ABBREVIATIONS

AOX1	<i>Alcohol oxidase 1</i>
BA	1,3:1,4- β -glucotetraose A
BB	1,3:1,4- β -glucotetraose B
BBG	Barley β -glucan
BC	1,3:1,4- β -glucotetraose C
BLAST	Basic Local Alignment Search Tool
BMGY	Buffered glycerol complex medium
BMMY	Buffered glycerol complex medium
CAZy	Carbohydrate-active enzymes
cDNA	Complementary DNA
CESA	Cellulose synthase
CMC	Carboxymethyl cellulose
CSL	Cellulose synthase like
CT	1,4- β -cellotetraose
ddH ₂ O	Double distilled water
DNA	Deoxyribonucleic acid
DGM	Di-galactosyl mannopentaose
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EG16	Endoglucanase 16
ER	Endoplasmic reticulum
EtOH	Ethanol
GAXs	Glucuronoarabinoxylans
GHs	Glycoside hydrolases
GM	Galactosyl mannotriose
HEC	Hydroxyl-ethyl cellulose
HPLC	High pressure liquid chromatography
IPTG	Isopropyl β -D-1-thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
LB	Luria Bertani broth

LBA	Luria Bertani agar
LT	Laminaitetraose
MC	Methyl Cellulose
MeGlcA	Methyl-glucuronic acid
MLGs	Mixed linkage β -glucan
MXE	Mixed linkage β -glucan; xyloglucan endotransglycosylase
NCBI	National Center for Biotechnology Information
NEB	New England Biolabs
OD	Optical density
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PMSF	Phenylmethane sulfonyl fluoride
rpm	Rounds per minute
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SOC	Super optimal broth
TAE	Tris base/acetic acid/EDTA
TBE	Tris base/Borate/EDTA
TBS-T	Tris buffer saline tween 20
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
TXG	Tamarind seed xyloglucan
UPGMA	Unweighted pair group method with arithmetic mean
UV	Ultraviolet
XET	Xyloglucan endotransglycosylase
XyG	Xyloglucan
XTH	Xyloglucan endotransglycosylase/hydrolase
XGOs	Xyloglucan oligosaccharide
XT	1,4- β -D-xylotetraose
YPD	Yeast extract/peptone/dextrose medium
YPDS	Yeast extract/peptone/dextrose/sorbitol medium
V	Voltage

1. INTRODUCTION

1.1. *ARABIDOPSIS THALIANA*

Arabidopsis thaliana (Figure 1.1.), has been widely utilized as a model organism [1] in plant molecular biology, genetics, and genomics, and investigations of *A. thaliana* genetic alterations offer a framework for testing approaches and arranging investigations including other plants, particularly species with commercial importance. Having a shorter development time [2], small genome size [3], productivity, and the capacity to perform mutational screenings to saturation have all led to an increase in Arabidopsis research. Arabidopsis has the potential to deliver critical insights into how gene activity affects the development of commercial crops [2]. The Arabidopsis genome encodes 33 XTH genes [4] which are divided into six groups, and they differ by their substrate specificity and approaches to function [5]. Paralogous gene expression and regulation have evolved differentially to aid in the perpetuation of this huge gene family [6]. Thus, each gene is likely to play some specific role within the plant, and knowing how it works and what substrates it can act on will be essential in defining its role within the plant.



Figure 1.1 Parts of the flowering plant *Arabidopsis thaliana* [7]

On the other hand, For scientific research and industrial use, a thorough insight of the *Arabidopsis thaliana* genome, including a deep and precise comprehension of the expression of its genes involved, is hence critical. This objective is dependent on the advancement of recent genomic technologies, along with the information collecting and processing methods. Applying high-throughput transcriptome sequencing [8], a genome-wide investigation of *A. thaliana* gene expression profiles over multiple tissues [6] and broadening processes could be provided [9].

1.2. *EQUISETUM TELMATEIA*

The great horsetail, *Equisetum telmateia* (Figure 1.2.), is a member of Equisetacea family and a herbaceous perennial plant that develops during early spring, and occur as a bright yellow-colored non-photosynthetic fertile stem that bears spores instead of producing seeds. Late spring is when sterile and photosynthetic green stems are developed, and late autumn is when they decay. These plants are lives in moist places that are typically nearby rivers and streams in Europe, most parts of Africa, and North America [10].



Figure 1.2. The great horsetail, *Equisetum telmateia* [47]

Stems of the great horsetail, from the fern genus, frequently mimic the stems of rushes and often seem to be comprised of several smaller components. Stems are porous and silica-reinforced and can develop as erect, upright shoots or creep along the surface. Horsetail

stems are made robust and toughened by silica [11]. Even though several of these species frequently hybridize with each other, the genus *Equisetum* contains a total of 15 species. *Equisetum* is classified within the Pteridophyta group, and they belong to the class Equisetopsida of ferns. Traditional medicine has used the aerial sections [12] of *Equisetum telmateia* as a supply of biologically active substances that cure inflammatory, diarrheal, abdomen, dermatological, and oral diseases.

The *Equisetum* structure is distinctive among living plants (Figure 1.3.). The stem is comprised of fragile, easily detachable internodes that are silica-rich and develop toward the base of the plant, just like in the Poales. Instead of starting in the axils like all other plants, lateral shoots in a whorl start between the leaves. At the same time, genome sequencing for transcriptome analysis on 108 (2n=216) chromosomes of all *Equisetum* species is difficult to complete.

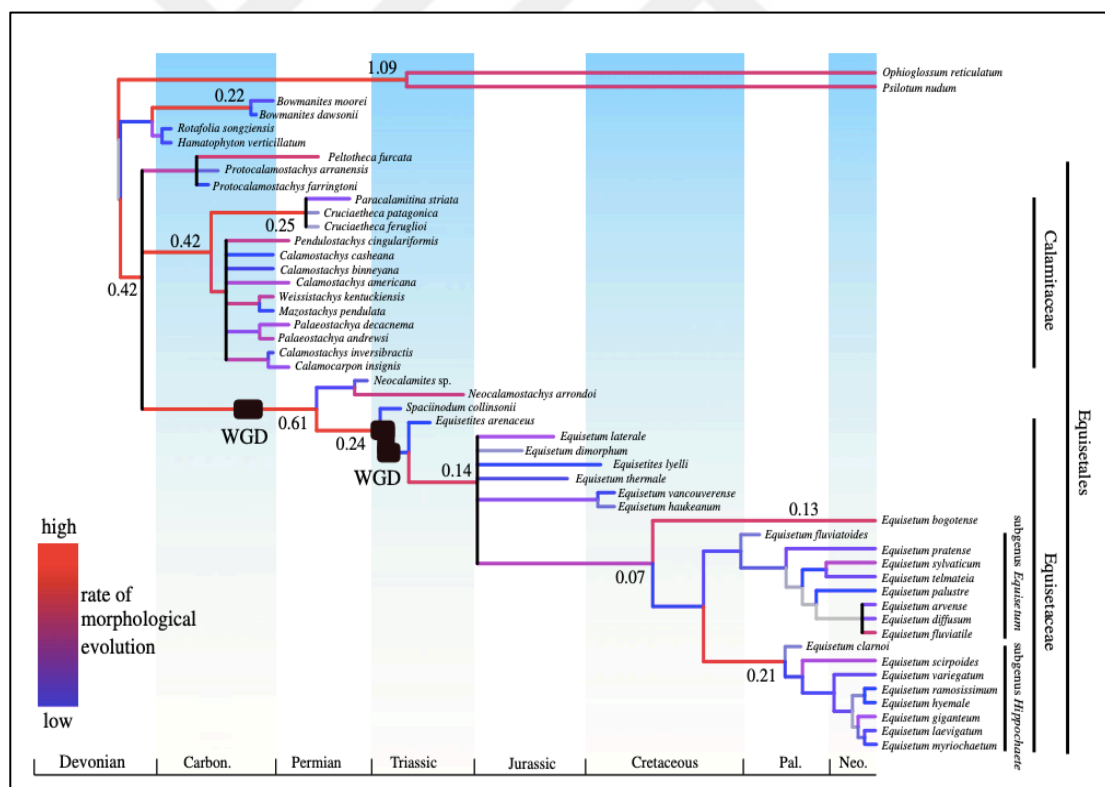


Figure 1.3. Phylogenetic tree of ancient and present Equisetales. The diagram depicts the position of each possible WGD (whole genome duplication). Phenotypic evolution rate indicated as low rates in blue and high rates are in red. [75]

1.3. PLANT CELL WALL

The primary sources of plant biomass are the division and elongation of cells. The most abundant biopolymer in the world, cellulose, is a key element of primary cell walls. De novo formation during cytokinesis and integration of incoming structural components into the existing cell wall are the two main steps of cell wall development. The cytoskeleton comes into contact with cell wall components. Approximately ten percent of certain genetic materials of plants are specialized in cell wall construction and remodeling, with dozens of additional genes participating in cytoskeleton and particle transportation [13]. Land plants have a distinctive property which primarily consist of carbohydrates in the plant cell wall. Plant growth and development are conducted by plant cells.

In plants, there are approximately 40 different types of cells. Differentiation of these cells could be determined by the biochemistry, morphology, and structure of the cell wall. Cellular differentiation can be seen in the same species as well, caused by the environmental factors and development stages. In early developmental stages, positive pressure in plant xylem [14] is inhibited by turgor pressure that occurred within the cells and ensures enlargement by an osmotic flow of the water, enabling the cell wall to maintain its shape and stand upright. Plant cell walls have improved defense mechanisms to conserve them against abiotic stresses and pathogenic microorganisms. Plant cell walls could be classified into two categories: primary and secondary walls. The secondary cell wall is maintained in certain cells once cell development and elongation have ceased, whereas the primary cell wall is generated during cell growth. Primary cell walls are extendable and flexible, because since they are rigid structures that are destroyed outside the plasma membrane. They don't get in the way of cell expansion and adjust to the plant's physiological requirements.

Plant cells are enclosed by cell walls as an essential characteristic and a major source of cellulose. Cell-cell adhesion and intercellular communication are controlled by the selectively permeable cell wall. All the vascular plant cells have a distinguishable plant cell walls. The cell wall is an instance of an extracellular matrix. During the development, plant cell walls mostly do not move, but differentiate with respect to their construction and function. Development of larger plant structure begins from the embryo or meristem, regions outside of the plasma membrane of single cells comprise of tightly connected

primary walls [15]. The plant cell wall, which organize a form of the plant by protecting and interacting with the plant cells, and support growth and development [16], and is a kind of extracellular matrix. Plant cells were enclosed by polysaccharide-rich walls that prevented them from bursting while remaining flexible enough to handle cell expansion. The cell walls of dicotyledon plants contain cellulose microfibrils. Selected one species for this thesis, *Arabidopsis thaliana* is a model for dicotyledons.

The cell wall changes during cell division, cell expansion, and differentiation, and the width of the cell wall range between 0.1 to 1 μm . Cell walls strengthen the cells despite their thin and flexible composition. *In silico* studies at the genomic and proteomic levels of XTHs from different plant species showed that distinct elements of contemporary XTHs evolved before of vascular plant diversification. In vascular plants, cell elongation correlated with the XET activity herewith highlights the inherent role of XTH in the evolutionary development of elongation and re-structuring of the cell wall [17].

1.4. PLANT CELL WALL STRUCTURE

Plants consist of organs; roots stems and leaves, tissues; ground, vascular and dermal, and, cells. Vascular plants (Tracheophyta) include horsetails, angiosperms, etc. that comprise vascular tissue such as the xylem, phloem, and roots, leaves. Primary cell wall has a thin structure which presents in the extracellular matrix of the plasma membrane, which could be seen during the growth and differentiation of the cells. When the cell enlargement is completed, the secondary cell wall forms consist of mostly lignin, which supplies mechanical strength by dictating the shape and size of the cells, and defense against pathogens including its thick structure.

The xyloglucan endotransglucosylase/hydrolase (XTH) family belonging to plant-specific extracellular enzymes, could easily split and reconnect xyloglucan chains [18] except for the requirement of that active donors. According to this unique feature, these enzymes are believed to be involved in the cellulose-xyloglucan network's post-synthetic remodeling during plant development. The deep and complex evolutionary history of *Equisetum*, originated from the Sphenopsida class with species richness, and traced back to Devonian [19].

1.5. CELLULOSE

Cellulose is a linear polysaccharide produced by plants, composed of repeated glucose residues that contain β -(1,4)-linkages both parallel and linear, forming long and rigid microfibrils which serve as the structure's load-bearing components. However, the scaffold of the cell wall is obtained by microfibrils connecting pectin and hemicelluloses [20]. To put it all in simple terms, the plant cell wall is an involved and active system [21] containing polymers and structural proteins which play a crucial part in the life cycle of the plant. It contributes to the growth of the plant, cell differentiation, defense, movement of the water, and communication intracellularly. Cellulose, which is the biomass of the plant, is found as insoluble microfibrils [22]. Cellulose microfibrils are found in soluble polysaccharides. Among these polysaccharides, cellulose and xyloglucan have similar β -1,4-glucan structures except for the water solubility of xyloglucan through the side chain modifications. Cell wall strength is maintained by cellulose microfibrils. Glucans in crystalline domains are organized, and cellulose microfibrils are mainly crystalline. Non-crystalline sections are also present in small quantities. Hydrogen bonds are degraded in non-crystalline regions, and the regular structure is lost. Due to the twist and torsions, the glucose chains are not parallel and linear. Crystalline regions make up 40 to 95 percent of cellulose, whereas non-crystalline regions make up the balance.

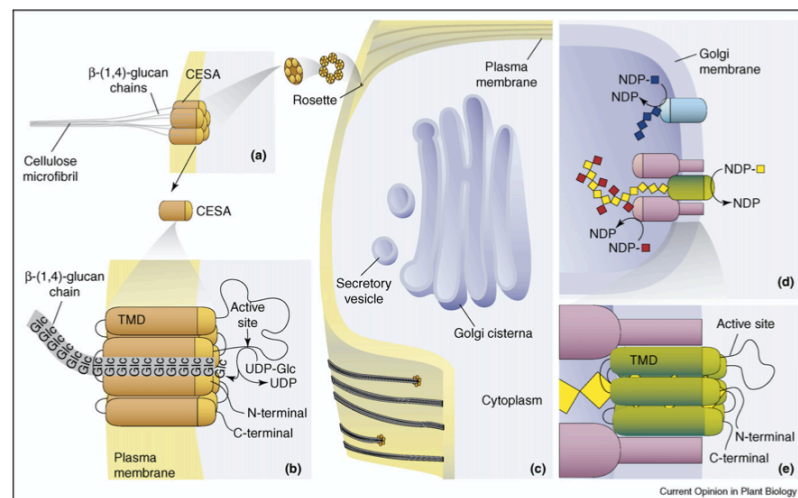


Figure 1.4. Cell wall polysaccharide biosynthesis [23]

The structural network formed by xyloglucan and non-covalently bound cellulose has a function in establishing the physicochemical characteristics of the cell wall, which is composed of eighty to ninety percent of carbohydrates. At the cell membrane, certain proteins are employed to polymerize and crystallize glucan chains, known as plant cellulose synthase (CESA) proteins. The expanding glucan chain, which cohered to form cellulose microfibrils, receives glucose residue from cytosolic uridine diphosphate-glucose (UDP-glucose) through the action of CESA enzymes. At the same time, the catalytic subunit is strictly connected to the expanding end of glucans (Figure 1.4.).

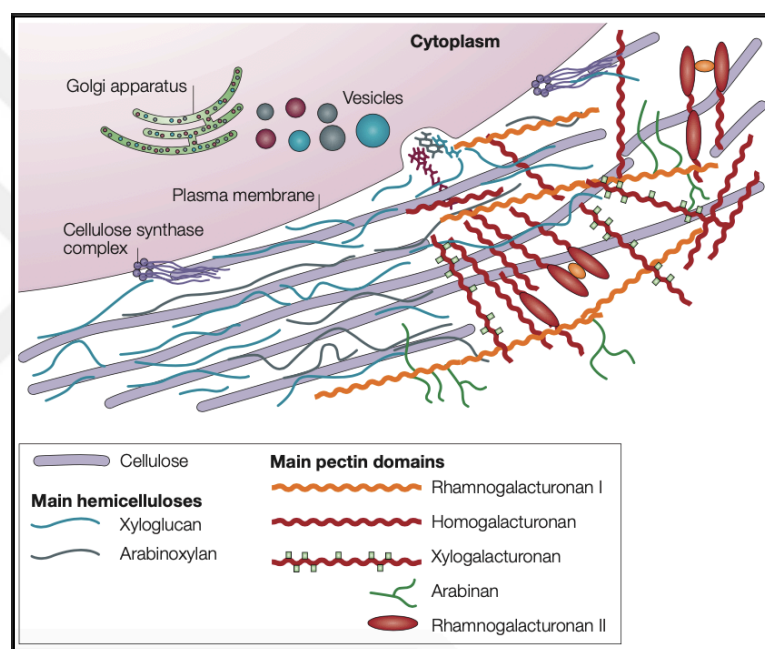


Figure 1.5. Primary cell wall structure [24]

Plant cell wall (PCW) polysaccharides are being converted into renewable energy sources and used to give nutritional advantages. A variety of alterations and modifications to the wall components are required to improve plant performance. It is difficult to make major modifications in the cellulosic process, according to research. The impact of genetically modified and non-GM technologies on human health and biofuels by altering cell wall structure is still being studied. In addition, the cell wall protects the cell against pathogens.

Extracellular components that interact with cell wall elements in healthy plants activate the cell's defense mechanism. The presence of signal transduction chains that can be activated by glycans is also required for this process to function. It is assumed that complex

chemicals are generated during cell wall breakdown to activate the immune system. Cell wall polysaccharide fragments and proteoglycans have been included in other investigations. Both primary and secondary cell walls are mostly found in plants with height. The primary cell wall is a bendable, thin, and highly hydrated layer that encloses the growing cell, whereas the secondary cell wall is a thicker, more rigid structure that begins to deposit once the cell stops growing. In short, primary cell walls are extendable, secondary cell walls are non-extendable. Plants are rich in carbohydrates known as cellulose, which is mostly expressed in vascular plants, although they may also be found in algae, bacteria, etc. The number of glucans in primary cell walls can range from hundreds to thousands, whereas secondary cell walls can have up to more than ten thousand glucans. Glucose chains include hydrogen bonds in between, and glucose chains aggregate because of the presence of intermolecular forces including hydrogen bonds and van der Waals forces, which cause crystalline cellulose that is found in two forms in nature [25]. When cellulose crystallizes, it becomes insoluble.

Throughout cell growth, development, and differentiation, the cell wall is made up of a variety of complex and varied structures. Cell-cell adhesion and intercellular communication are controlled by the selectively permeable cell wall. In a plant, there are around 40 distinct types of cells. The chemistry, size, format, and components of the cell wall can all be used to differentiate them. The fact that even plants of the same species exhibit cellular differentiation is due to environmental factors and plant developmental phases.

1.6. XYLOGLUCAN

The cell wall is a rough and flexible structure with cellulosic microfibrils and polymers formed by secreted polysaccharides of the extracellular matrix. The expansive growth of the cell wall has not been completely explained since the interactions between the polysaccharides are assumed to be generated by physicochemical interactions and enzymatic reactions.

Xyloglucan (XyG) is a basic cross-linking molecule that could be identified as a greatly expressed hemicellulose and have a feature in cell degradation and proliferation in dicotyledonous plants [26] depending on the ability of cellulose molecules to attach and

The attachment of XyG and xylosyl-residue promotes cell stability between these microfibrils. The side chain xylosyl groups also may be replaced using certain saccharides, the arrangement of these substitutions results in the wide variety of XyG structural motifs identified in various species of plants. Furthermore, XyG and pectin have covalently links in between. [29] Expansins, diverse glycosyl hydrolases, and XyG endotransglycosylases are only several of the apoplastic proteins that have been revealed to interact with XyG. Thus, it is assumed that the metabolism of XyG and turnover are crucial for cell elongation.

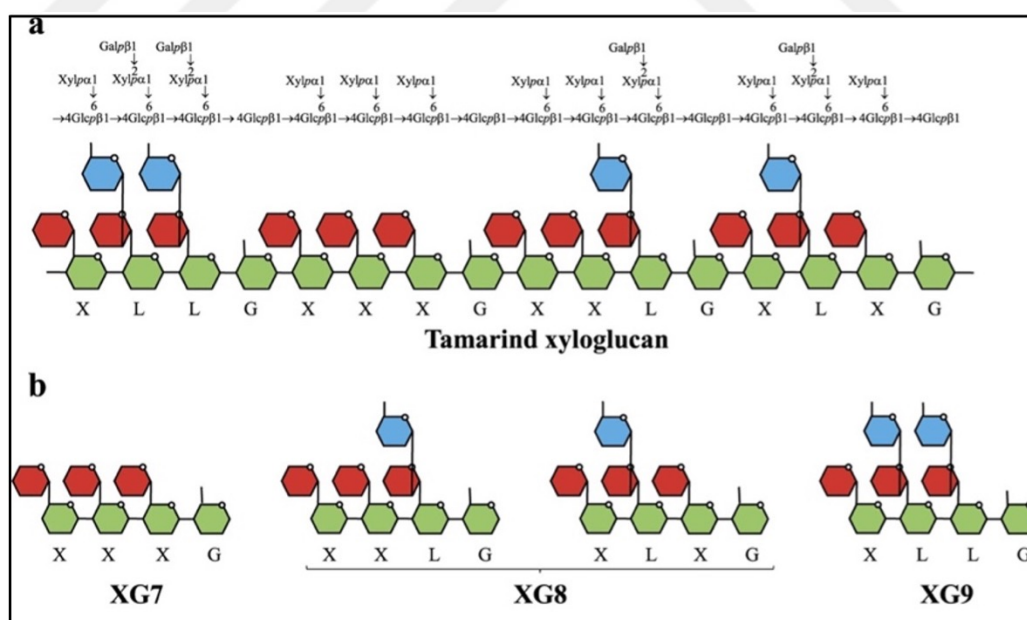


Figure 1.6. Tamarind xyloglucan, XG7, XG8, and XG9 schematic representation. a) Tamarind xyloglucan general structure b) Glucosyl xyloglucan oligosaccharides (XGO₄)

A plant absent XyG would exhibit significant cellular and tissue abnormalities, according to this function. XyG is not as crucial for wall construction and activity as believed earlier, as shown by the comparatively minor growth phenotypes of an *Arabidopsis* mutant missing observable XyG [31]. The quantity of xyloglucan could change depending on the species; XyG is mostly abundant in flowering plants in the primary cell walls, and also found in secondary cell wall with the robust and non-lignified forms of cotyledons or seed endosperms [32]. The interaction of xylosylated glucan sidechains with xylosyl residues results in sidechain variety and structural nomenclature, which is defined as one letter nomenclature [28]. A one-letter code was developed for each of the XyG sidechains identified in Figure 1.7. The backbone $-(1,4)$ -glucosyl residues are marked in black. An underline denotes acetylated sidechains, that are in L or S. The sidechains of xyloglucans are generally beginning with $-D$ -Xylp residues, and backbone attached to arabinose to form arabinoxylan which found in walls of cereals and grasses [33], or attached to glucuronic acid and form glucuronoxylan, which is essential for the secondary wall of dicotyledon plants [34].

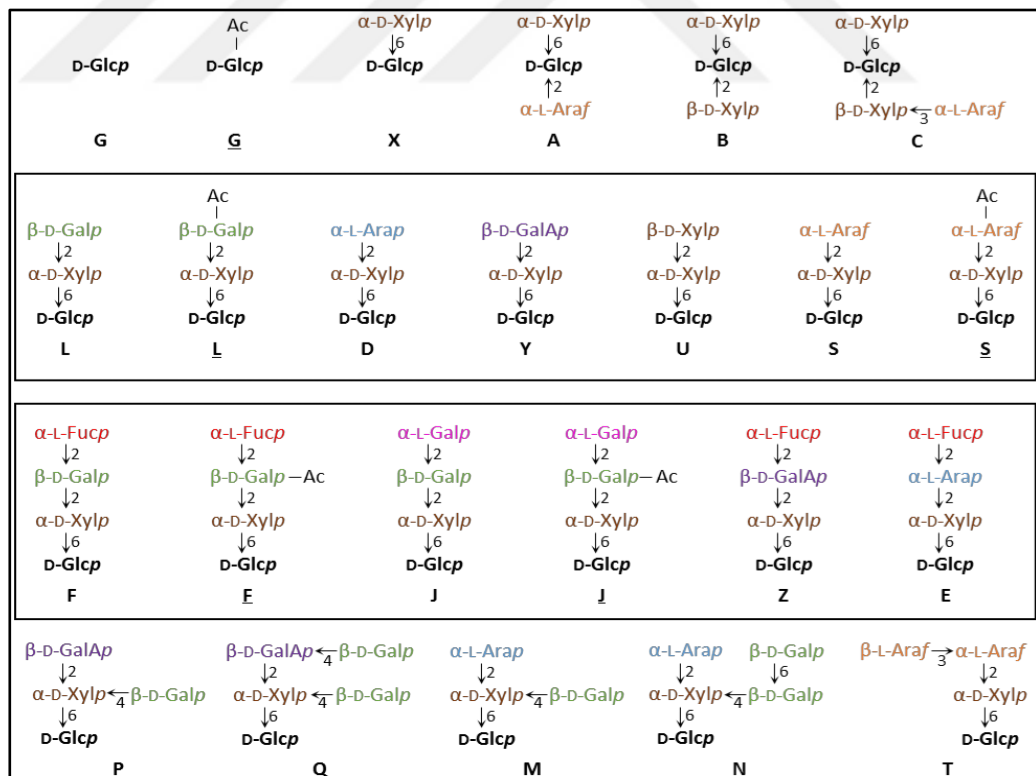


Figure 1.7. A broad range of XyG sidechains [28].

Equisetaceae family, have XXXG xylosylation motif with X, L, F, D, E, and S sidechains whereas *Arabidopsis thaliana* from Brassicaceae family, galactosyl substituents are substituted to the O-acetyl residue located in the O-6 position. The length of the XG backbone is crucial for conducting the interactions, still, specification of an optimal length for the XG backbone has been debated.

All vascular plants that have been studied so far have primary walls which contain arabinoxyloglucan, fucogalactoxyloglucan, and arabinogalactoxyloglucans. These materials might comprise twenty percent dry weight of the cell wall. The labeled composite product has a large molecular weight that is different from the donor and acceptor substrates forms when the XTH enzyme reacts with the donor polysaccharide and acceptor substrate. An elevated test for XTH activity also employs fluorescently labeled acceptors with pyridylamino, sulphorhodamine, or fluorescein [32]. In a method referred to as N-glycosylation, an oligosaccharide and a carbohydrate made up of multiple sugar molecules, also known as a glycan, is linked to a nitrogen atom. The construction and activity of numerous eukaryotic proteins depend on this kind of linkage. The prevalence of xyloglucans in first land plants indicates that they played a key role in the change from moist to dry environments [35]. Products of XTH genes are essential therefore maintain in higher plants as members of XTH gene families connected with tissue or time-dependent conditions [32].

1.7. MIXED-LINKAGE B-GLUCANS

Mixed-linkage β -glucans (MLGs) is unsubstituted linear polysaccharides structured with a (1;3,1;4)- β -D-glucopyranose backbone, and are generally related to *Equisetales* (*Equisetum* species), algae, and Poaceae family that include species of grasses, grains, and cereals. Biosynthesis for MLGs appears to be significantly maintained or to have formed again through convergent evolution because it occurs in two genetically highly different land-plant taxa [36]. They are intensively found in the endosperm walls of signified species above, which are widely eaten up for daily intake, making them a very essential element of human nutrition and a vital molecule associated with conserving the glucose in plant cell walls.

Cell walls are made of xyloglucan, which is transglycosylated from MLG by the enzyme xyloglucan endotransglucosylase (MXE) [36]. Equisetum species yielded tetra saccharide repeat-unit (G4G4G3G). MLGs are comprised of (1;3,1;4)- β - D-glucopyranose backbone with asymmetrically attached (1,4)- β -D-glucosyl residues within at least two or greater quantities of (1,4)- β -D-glucosyl residues. The (1;3,1;4)- β -D-glucosyl residues lack of sequences that have constant repeating. This asymmetry of polysaccharides comes from the kinks in (1,3)- β -D- glucosyl residues which ensure polysaccharides become soluble due to polymerization.

1.8. XYLAN

Xylans are natural polymers that have arabinose, glucuronic acid, or 4-O-methyl-glucuronic acid replacement at the O-2 or O-3 positions, respectively, on their (1,4)- β -D-linked xylose backbone. However they may also be found in primary cell walls of grains and other grass species, they are typically found in secondary cell walls. In dicots, xylans make up about 5 percent of the polysaccharide, but they make up 30 percent of the polysaccharide in primary walls. They're a type of polysaccharide found in the secondary cell wall. Arabinoxylan and glucuronoarabinoxylan are two types of xylan. Backbone composition differs between species. The biosynthesis of xylan was aided by a variety of enzymes. Two distinct protein families, GT43 [37] and GT47 [38] are engaged in xylan synthesis, particularly in *Arabidopsis thaliana*. These proteins contribute to the xylan chain or backbone elongation [39]. Xylan has been considered a source for the manufacture of biofuels since it contains polysaccharides including pentoses [40]. Since several enzymes would be required to break down such pentoses, fermentation would be challenging.

1.9. MANNAN

Mannans are one of the taxonomically extensive polysaccharides that comprise of β -1,4-mannose backbone or a mixture of mannose residues with β -1,4- linked glucose. Green algae and grassland plants are the most common hosts. Pure mannan, glucomannan, galactomannan, and galactoglucomannans are all examples of other polysaccharides. The backbone belonging to glucomannan is mostly composed of glucose and mannose residues.

Solubility and contact with several other carbohydrates are enhanced by galactosyl residues. In vascular plants, such as in model angiosperm *Arabidopsis thaliana*, mannans are present in secondary cell walls of xylem tissues. Galactomannans are ubiquitous in tissues that take part in the storage of the wall components, particularly in seed endosperm of the Fabaceae family [41]. Despite their presence in the seeds of various plants, their roles still are undetermined. They are found in the *Arabidopsis thaliana* cell wall. Furthermore, despite the removal of the mannan synthase genes, no morphological alteration has been detected [42].

1.10. PECTIN

Pectins are a class of physically complicated polysaccharides with a high concentration of galacturonic acid residues. In higher plants and gymnosperms' expanding cells, pectin levels are highly ubiquitous. They are typically present within the soft regions of plants including the middle lamella and xylem. In dicotyledons and monocotyledons which lack ferulic acid in their cell walls known as non-commelinid, pectin makes up about 35 percent of the major cell walls. Pectin is indeed synthesized in the Golgi and transported to the wall by the Golgi apparatus [43]. The key element seems to be that knowledge of pectin synthesis and composition is essential for comprehending and manipulating cell walls in plants. Numerous researchers [44] have examined it because of its industrial significance. Numerous significant roles, including those in growth and expansion, resistance, morphology, and cell wall permeability, have been identified. It has been employed as a stabilization ingredient in food and pharmaceuticals on an industrial scale. Additionally, there are several other studies relating to drug levels, such as lowering blood LDL levels and tumor progression [45].

1.11. XYLOGLUCAN ENDOTRANSGLYCOSYLASE/HYDROLASE

Cell wall rigidity, mobility, and permeability determine how land plants develop. During the cell expansion and differentiation processes, the cell wall alters, allowing cells to adapt to different functional requirements as well as cell wall stresses generated by the surroundings and pathogens. Water, nutrients, and regulators that affect the growth could

transfer through the cell wall, which is important for cell-to-cell interaction. Modifications toward the cell wall are needed during cell expansion and loosening. The wall polysaccharides regulate the width and density of cellulose microfibrils throughout the cell wall elongation process driven by turgor pressure. Xyloglucan – cellulose network modify cross-links of xyloglucans induce cell wall weakening to limit unmanageable cell growth.

The covalent bonds between various xyloglucan chains, including between xyloglucan and other non-cellulosic polysaccharides, are processed by xyloglucan endotransglycosylase/hydrolases (XTHs). The primary two catalytic activities of XTHs are the xyloglucan endo-transglycosylase function or integrated xyloglucan endotransglycosylase and xyloglucan endohydrolase function. Xyloglucan endohydrolases hydrolyze the (1,4)-glucosyl linkages belonging to the backbone of xyloglucan, however, xyloglucan endotransglycosylases break the (1,4)-glucan backbone and subsequently carry the nonreducing component onto some other molecules xyloglucan oligosaccharide. Endohydrolase activity could cleave the xyloglucan backbone as a section of the loosening mechanism, but a deficit of wall synthesis and strengthening causes a decrease in tensile strength.

Encoded proteins of XTH genes responsible for primarily two actions that have dissimilar influences on xyloglucan backbones are endotransglycosylase and endohydrolase activities.

Xyloglucan endotransglycosylases (XET) nonhydrolytically cleave and ligate the (1,4)- β -D-glucosyl linkages while xyloglucan endohydrolase is responsible for the hydrolyzation of xyloglucan backbones permanently [46]. Accordingly, endotransglycosylase activity is required for cell wall stability due to the cleavage of xyloglucan and subsequently switches to a different substrate. This can also strengthen the cell wall by merging newly produced xyloglucan. The power needed for the production of new glycosidic connections arises mostly from the breakdown of the existing xyloglucan backbone's glycosidic linkages. The glycoside hydrolase family of plant and microbial enzymes breakdown xyloglucan, however, these activities are identified in six GH families of microbial enzymes whereas plant enzymes are only identified in GH16 [46], contradicts the significance of this enzyme in the matter of xyloglucan cleavage.

XTH enzymes are divided into six clades based on their amino acid similarities: Group I, Group II, Group III-A, Group III-B, Group IV, and EG16 [47]. The molecular mechanisms of the XTH enzymes do not differ between groups, although Groups I and II possess XET activity while Group III possesses XEH activity [48].

1.12. POLYSACCHARIDE AND OLIGOSACCHARIDE SUBSTRATES

The byzantine macromolecular structure [49], plant cell wall, mostly made up of carbohydrates; cellulose, hemicellulose and pectin, and glycoproteins and lignin, distinctly from the animal extracellular matrix by the composition which is mostly composed of nitrogen-containing polymers. Living organisms convert inorganic carbon into organic molecules by photosynthesis, a process called carbon assimilation which is essential to utilize energy storage by fixing the carbon. For this reason, polysaccharides within the plant cell wall are a renewable carbon source.

Further, polysaccharide-rich plant biomass is one of the most important sources of carbon, because the cell wall accounts for the majority of plant biomass. Plant polysaccharides are studied for a long time to get a better knowledge of plant growth and development, identifying interactions of polymers, biomass synthesis, and biofuels.

1.13. HETEROLOGOUS EXPRESSION AND THE PICHIA PASTORIS SYSTEM

Plant structure and biochemical alterations require to be properly investigated to be capable of adapting plants to generate plant-based products e.g. food, fiber, papermaking, brewing, and textile. Besides, it's challenging to investigate a living organism. The invention of biotechnologically or pharmaceutically important recombinant plasmids and the bacterial transformation that enables the recombinant protein expression represents a significant advancement. Eukaryotic protein expressions typically lead to the formation of inclusion bodies for the reason that the absence of post-translational modifications for instance; N-linked glycosylation and disulfide bond establishment in secretory expression. Later on, the *S. cerevisiae* has been widely used as a platform that has limitations like hyperglycosylation and decreased yields [50].

The methylotrophic yeast, *Pichia pastoris*, has conveniences of using over other platforms, which as through secretory expression, recombinant proteins can be synthesized both intracellularly and extracellularly and higher cell density. *P. pastoris* uses the Methanol Utilization Pathway (MUT) to build biomass and generate energy by metabolizing methanol as a carbon source. This promising yeast has two AOX genes, AOX1 and AOX2, which express two alcohol oxidase enzymes. AOX1 generates the primary enzyme that accounts for up to thirty percent of all soluble proteins in yeast induced with methanol, demonstrating AOX1's high strength of promoter *P. pastoris* [50]. The operational investigation of plant proteomes remains to be a significant obstacle in plant gene expression heterologously, in particular genes expressing carbohydrate-active enzymes (CAZymes) including glycosyltransferases and glycoside hydrolases. There are currently several practical approaches for expressing recombinant plant enzymes which are active and in sufficient quantities for characterization in terms of physio-chemically and structurally [51]. Robust transgenic lines are produced by homologous recombination, which integrates linearized foreign DNA by restriction endonucleases into the host genome. Additionally, *P. pastoris* is reported to produce small amounts of endogenous proteins that aid in the downstream purifying process.

Numerous steps of heterologous protein expression in *Pichia pastoris* platform should be examined. The DNA sequence of the expression target is cloned into an expression vector consisting of a convenient promoter, selection marker gene to pick out the transformant cells, and sequence of transcription termination which is encoded by a multiple cloning site (MCS). To maintain stability, a vector containing the protein-coding DNA sequence is transformed into yeast by the electroporation method. The target protein is then heterologously expressed in optimal culture conditions, and the pure and condensed protein is produced which is a must for enzyme characterization. *P. pastoris* is a convenient host for recombinant protein production by the heterologous expression, although it has some certain limitations such as protein degradation during secretion. By adjusting the pH and temperature along with utilizing protease deficient yeast strains as hosts, appropriate culture parameters could be supplied to avoid degradation. Several of the most significant reasons to use *P. pastoris* as a host system is its qualification to N-linked glycosylation which is required for stability and optimal performance. For this reason, the selected host system ought to be capable of glycosylation of the protein after translation which target

enzymes have also glycosylated. Later in the translation, signal peptidase eliminates the signal sequence which could be predicted by SignalP 6.0 software program and arise in the mature protein.

The host strain *Pichia pastoris* SMD1168H lacks vacuole peptidase A (pep4), whereas this enzyme is crucial for carboxypeptidase Y and protease B1 activation, it also lacks foregoing proteolytic enzymes, and is selected for heterologous expression. To achieve optimum conditions, which may differ between multiple strains and externally produced proteins; concentration of methanol, Mut types, temperature, and incubation time should be modified. [52] Consequently, the target protein determines the optimum conditions for target protein production in the *P. pastoris* expression system.

Altogether, several advantages of protein expression by the *P. pastoris* method comprise correct folding in the ER and recombinant protein expression to the cell's external environment. The restricted synthesis of endogenous secretory proteins in the *P. pastoris* expression system further facilitates the purification of recombinant proteins. Lastly, the expression vector pPICZ α -C for *Pichia pastoris* is intended for rapid cloning and direct selection of integrant strains by containing the Zeocin™ resistance gene, high-level expression, detecting and purifying the recombinant protein quickly. Target DNA comprising XTH gene from *Arabidopsis thaliana* AtXTH30, and XTH gene from *Equisetum telmateia* EtXTH1 cloned into *P. pastoris* vector by use of T4 DNA Ligase. Plasmid maps of cloned target DNA into the expression vector pPICZ α -C (Figure 1.8).

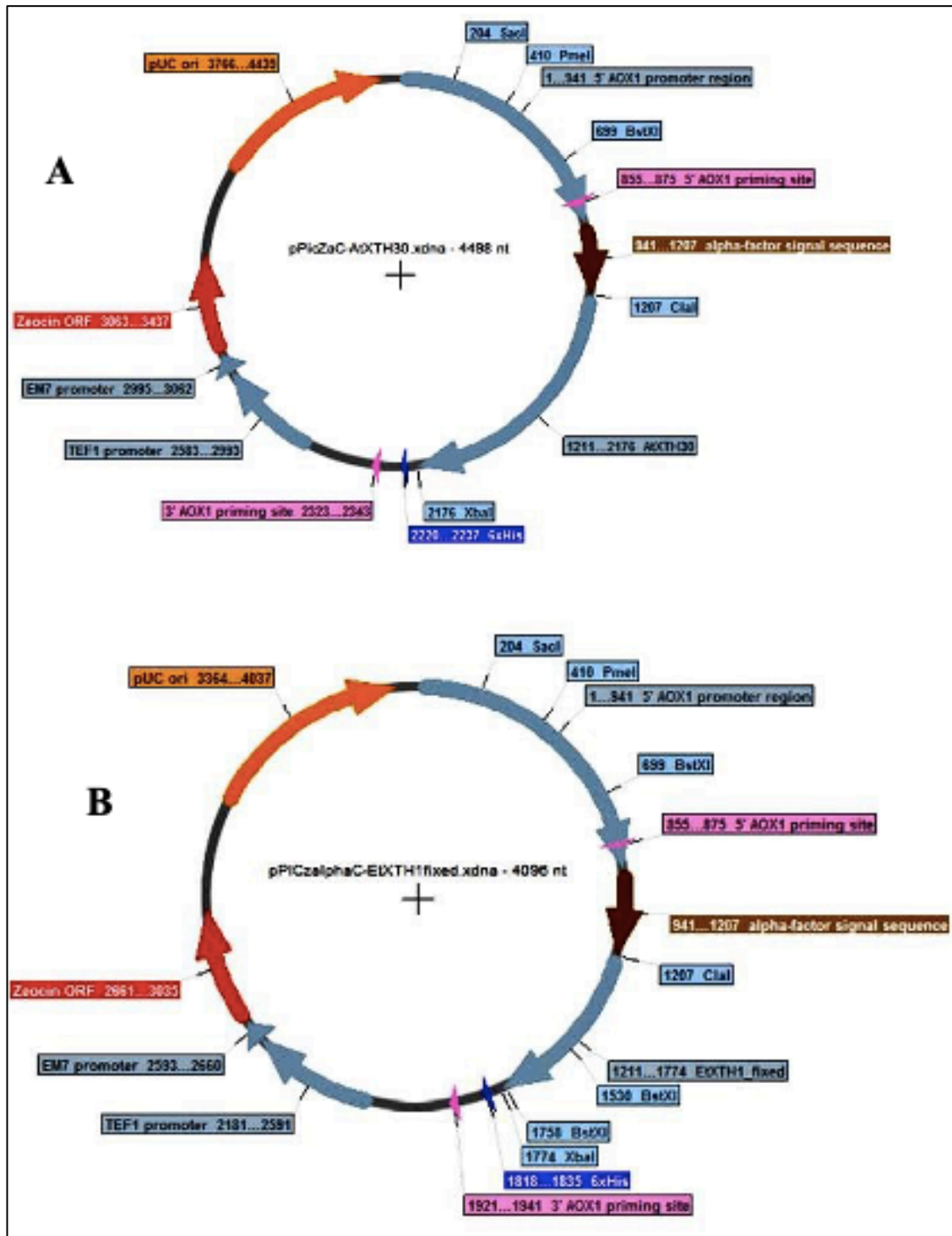


Figure 1.8. Plasmid map of target DNA cloned into expression vector pPICZα-C. Zeocin highlighted in red. A) *AtXTH30* and the expression vector have been cut from ClaI and XbaI restriction sites. For linearization, plasmid has been cut from the SacI restriction site. B) *EtXTH1* plasmid has been cut from XbaI and ClaI restriction sites.

1.14. XYLOGLUCAN ENDOTRANSGLYCOSYLASE/HYDROLASE ENZYME CLASSIFICATION AND EVOLUTIONARY LINKAGES

Based on similarities in their amino acid sequences, XTH enzymes are divided into various subfamilies: Group I, Group II, Group III-A, Group III-B, Group IV, and EG16. It has been demonstrated that Group I, Group II, and Group III-B enzymes primarily have XET activity, while the Group III-A enzymes show XEH activity, even though that there is no difference among XTH enzymes being the part of different groups depending on their biochemical properties or mechanism of action. Groups, when they are completely investigated, could appear to provide an organ-specific expression profile. The GH16 family, which includes several enzymes with various catalytic functions, contains XTH enzymes.

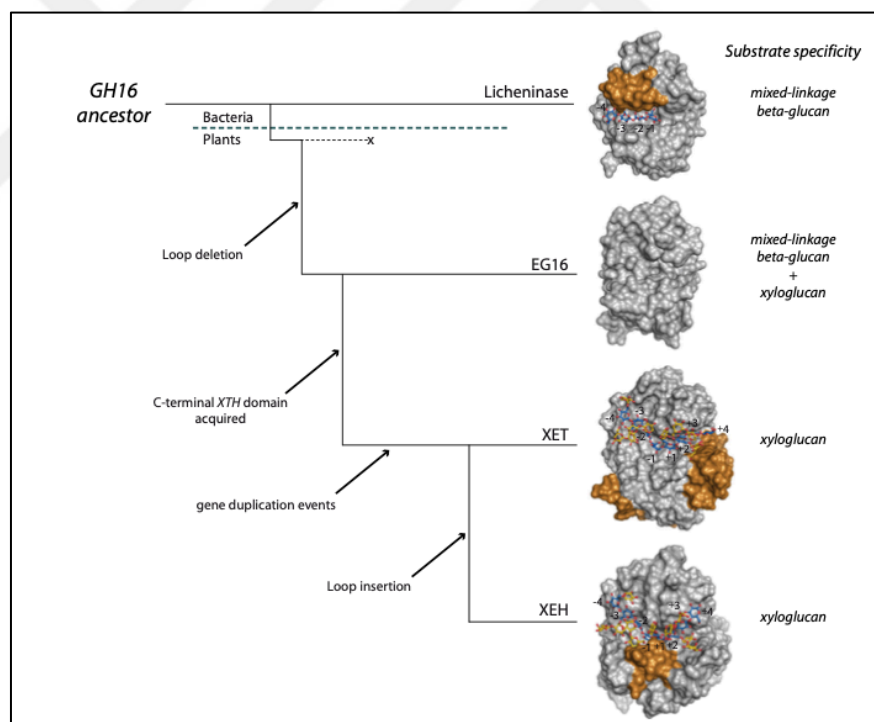


Figure 1.9. Evolution of EG16 licheninase-like ancestor XTH gene products. The gold-colored part indicates the loop or C-terminal extension respectively. [53]

Proteins of bacterial licheninases (Figure 1.9.) and XTHs were determined to become the proximate relatives inside the family, with both sequences bearing the 'DEIDFEFLG' catalytic region. All this and bacterial licheninases possess the β jelly roll fold, which is a

characteristic of the GH16 family. Comparing protein models of bacterial licheninases and XTH family members, significant structural component variations are apparent in the structures. It is likely to see the structural changes that have occurred during XTH using the relevant sequencing data and protein model datasets. It is likely to detect the structural changes that have occurred during the evolution of XTH by using the relevant sequencing data and protein model datasets.

Conserved Cysteine (Cys) residues located after the active site, which are predicted to contribute to stabilizing the C-terminal region, are another characteristic commonly observed in XTH enzymes. An N-linked glycosylation site after the active site is maintained by all XTH enzymes, except for a few Group III-A members; this site is regarded to be crucial for the stability of the protein and its enzymatic characteristics [53,54]. XTH clades additionally differ in loops 1 and 2, which are typically born insertions by representatives of Group III-A and Group III-B [53].

The clade in the XTH phylogenetic tree that has individuals with different active site amino acids is regarded as the ancestral clade, which was stated as Group IV newly. Group IV participants were divided into several groups as a result of this distinction and their resemblance to β -(1,3);(1,4)-glucanase ancestors in comparison to other clade members. In the evolution of XTH's from their ancestors, this group was assumed to be a connecting step. Due to this knowledge, a greater range of substrate specificity for this clade is predicted. Studying and comprehending the evolution of a cell wall-modifying enzyme highly depends on the enzyme characterization of members of the ancestral group.

1.15. AIM OF THE STUDY

The cell wall is a significant element of the plant since it defends and develops with its surroundings. Understanding the form and composition of the elements found inside the plant body is critical for preserving the plant from challenges and increasing yield. To achieve an understanding of modifications of the plant cell wall, the β -linked polysaccharide xyloglucan should be investigated which has important roles such as creating load-bearing structures in the plant cell wall.

The primary problem that permits cell growth and proliferation is the alteration through the deposition and reorganization of leading wall carbohydrates located in the cell wall. The modifications are performed by XTHs, which accelerate the breakdown and reformation of xyloglucan. In the cell wall, XTH enzymes were shown to be able to transglycosylase various cellulosic and non-cellulosic carbohydrates, and they play a significant part in numerous cell wall functions such as growth, shortening, signaling, and in return for various stresses. With all of the features that XTH enzymes could have, characterization of the enzymatic activities and analysis of the roles of their gene products are very important to improve our knowledge of such enzymes. The objective of this thesis was to investigate *Arabidopsis* and *Equisetum* XTH enzymes at the genomic, transcriptome, and protein levels to gain an exceptional insight into how these enzymes work and what functions they serve. Understanding the roles of the enzymes that change the carbohydrate matrix is essential for altering cell wall characteristics to increase yield and preserve the plant from external factors.

Recent studies on other *Equisetum* species have revealed interesting, broad substrate specificity ranges for some of their XTH enzymes. More detailed analyses of *Equisetum* XTH enzyme action patterns and substrate preferences can help to elucidate the evolution of this important gene family. Focusing on genes that belong to the poorly characterized Group III-B and genes that express enzymes with expansive substrate specificities such as those from Group IV (Floral Group), is essential due to the lack of published studies and carrying out a study for selected genes in Group III-B to identify which polysaccharides these genes choose as our target. The findings of this study would be compared to model plant species such as *A. thaliana* and *O. sativa*, primarily, for understanding further specifics about the production of xyloglucan in *Equisetum* and the variety of genes participating.

2. MATERIALS

2.1. EQUIPMENT

The equipment was supplied from different companies and given in Table 2.1.

Table 2.1. List of equipment and company names.

Equipment	Company
AKTA Primer Plus Chromatography System	GE Healthcare Life Sciences, UK
New Brunswick 44/44R Incubator Shaker	Eppendorf Innova, USA
1100 Series High-Performance Liquid Chromatography (HPLC)	Agilent Technologies, USA
T100TM Thermal Cycler	Bio-rad, USA
UV Transilluminator	Sigma-Aldrich, USA
Electrophoresis Power Supply omniPAC CS-300V	Cleaver Scientific, UK
Rocking Platform Shaker	VWR, USA
SUB Aqua 26 Plus Unstirred Water Bath	Grant, UK
Eppendorf™ Centrifuge 5424 R	Eppendorf Innova, USA
Avanti J-E High Speed Centrifuge	Beckman Coulter, USA
ChemiDoc XRS+ System	Bio-rad, USA
Drying and Heating Chamber	Binder, Germany
Elite 300 Plus Power Supply	Wealtec, USA
Forma 900 Series -86 C Upright Ultra-Low Temperature Freezer	Thermo Fisher Scientific, USA
Gene Pulser	Bio-rad, USA
Heated/Refrigerated Circulator	VWR, USA
Micropipettes	Eppendorf Innova, USA
Mini Vertical Electrophoresis Unit	Thermo Fisher Scientific, USA
MS-H280-Pro Circular-top LED Digital Hotplate Stirrers	Scilogex, USA
BioSpec-Nano Small Volume Spectrophotometer	Shimadzu Corporation, Japan

2.2. CHEMICALS & DISPOSABLES

Chemicals and disposables were supplied from several companies according to Table 2.2.

Table 2.2. List of chemicals and disposables with company names supplied by

Chemicals & Disposables	Company
0.2 mL PCR Tubes	ISOLAB, Germany
1.5 and 2 mL Centrifuge Tube	ISOLAB, Germany
10, 200, 1000 µl Pipette Tips	Axygen Scientific, USA
Laboratory Film	Parafilm "M", USA
Petri Plates (100 mm x 15 mm)	ISOLAB, Germany
BioSep-SEC 4000 Column	Phenomenex, USA
HiPrep 26/10 Desalting Column	GE Healthcare Life Sciences, UK
His-Trap FF 5 ml Column	GE Healthcare Life Sciences, UK
Inoculation Loops	ISOLAB, Germany
Minisart [®] SRP15 Syringe Filters	Sartorius, Germany
Amersham [™] ECL [™] Prime Western Blotting Detection Reagent	GE Healthcare Life Sciences, UK
PCR Master Mix, 2X, K0171	Thermo Fisher Scientific, USA
Anti-6x His-tag Antibody	Abcam, UK
1 Kb Plus DNA Ladder	Thermo Fisher Scientific, USA
10X Green Restriction Buffer	Thermo Fisher Scientific, USA
10X Reaction Buffer	GeneOn, Germany
2-Propanol, Ammonium di-Hydrogen Phosphate	PanReac AppliChem, Spain
100 bp DNA Ladder	Thermo Fisher Scientific, USA
UltraPure TBE Buffer, 10X	Thermo Fisher Scientific, USA
6X DNA Loading Dye	Thermo Fisher Scientific, USA
Acetic acid 100%	Sigma Aldrich, USA
Acetone ≥ 99.5%	Sigma Aldrich, USA
Acrylamide 30%	Sigma Aldrich, USA

Agar	Sigma Aldrich, USA
Agarose	Sigma Aldrich, USA
Ammonium Phosphate Monobasic	Sigma Aldrich, USA
Ammonium acetate	Sigma Aldrich, USA
Calcium Chloride (CaCl ₂)	Avantor, USA
ClaI	Thermo Fisher Scientific, USA
Nuclease free water (DNase/RNase free)	Invitrogen, USA
Dithiothreitol (DTT)	Thermo Fisher Scientific, USA
D-sorbitol	PanReac AppliChem, Spain
Ethanol Absolute	Sigma Aldrich, USA
Ethidium Bromide	Sigma Aldrich, USA
Formic Acid 98-100%	Sigma Aldrich, USA
Nucleospin Gel and PCR Clean-Up Kit	Macherey-Nagel, Germany
Gene Pulser/Micro Pulser Electroporation Cuvette	Bio-Rad Laboratories, USA
Glycerol 99.0%	Sigma Aldrich, USA
Isopropyl-β-D-thiogalactopyranosid (IPTG)	PEQLAB Biotechnology, Germany
Immun-Star™ WesternC™ Chemiluminescent Kit	Bio-Rad Laboratories, USA
Kanamycin Sulphate	Sigma Aldrich, USA
Lyticase from <i>Arthrobacter Luteus</i> Lyophilized Powder	Sigma Aldrich, USA
SacI	Thermo Fisher Scientific, USA
Plasmid DNA MiniPrep Kit	Thermo Fisher Scientific, USA
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific, USA
Peptone	Merck Millipore, USA
pGEM T-Easy Vector System	Promega. USA
Phenol Solution	Sigma Aldrich, USA
Phenol:Chloroform 5:1	Sigma Aldrich, USA
Potassium Chloride ≥ 99.0%	Sigma Aldrich, USA
Potassium Phosphate Dibasic	Sigma Aldrich, USA
Potassium Phosphate Monobasic	Sigma Aldrich, USA

pPICZ α A, B, & C <i>Pichia</i> Expression Vectors	Thermo Fisher Scientific, USA
Sodium Acetate	Sigma Aldrich, USA
Sodium Carbonate anhydrous	PanReac AppliChem, Spain
Sodium Chloride	Sigma Aldrich, USA
Sodium Dodecyl Sulphate	PanReac AppliChem, Spain
D-(+)- Glucose	Sigma Aldrich, USA
Sodium Phosphate Monobasic Dihydrate	Sigma Aldrich, USA
T4 DNA Ligase	Thermo Fisher Scientific, USA
Tetramethylethylenediamine (TEMED)	PanReac AppliChem, Spain
Trichloroacetic Acid (TCA)	Merck Millipore, USA
Tris Buffered Saline (TBS) Powder	Santa Cruz Biotechnology, USA
XbaI	Thermo Fisher Scientific, USA
Yeast Extract	Sigma-Aldrich, USA
Yeast Nitrogen Base	Sigma-Aldrich, USA
Zeocin	Thermo Fisher Scientific, USA
X-Gal	Thermo Fisher Scientific, USA
Isopropyl β -D-Thiogalactopyranoside (IPTG)	Thermo Fisher Scientific, USA

2.3. BIOLOGICAL MATERIAL

Biological materials; bacterial strains and yeast strains are used for expression according to Table 2.3.

Table 2.3. List of biological materials with company names supplied by.

Biological Material	Company
<i>E.coli</i> DH5 α Strain	Thermo Fisher Scientific, USA
<i>Pichia Pastoris</i> SMD1168H Strain	Thermo Fisher Scientific, USA
<i>Komagataella pastoris</i> ATCC28485 Strain	ATCC, USA

2.3.1. Reverse and Forward Primers Of Expression Target

Polymerase chain reaction requires single-stranded DNA pieces around 18-30 bases to hybridize the template DNA. Guanine-Cytosine (GC) content and melting temperature T_m are crucial parameters together with the cycling parameters of denaturation, annealing, and extension stages that have to be optimized. Forward and reverse primer sequences for *EtXTH1* genes were supplied from SenteBioLab, Turkey.

Table 2.4. Forward and reverse primer sequences for *EtXTH1*

Gene names	Forward primer sequence
<i>EtXTH1_mut_5'</i>	ACCTATCGATGAAAGTGTCCCCTGTTGCCACAAGTT
	Reverse primer sequence
<i>EtXTH1_mut_3'</i>	GCCTTCTAGACCGGCCACCCTTCTGGAAGTACATGG

Table 2.5. Specifications of forward *EtXTH1* primer

Oligo No	191113-1-6	GC	36.2%	Tm(basic)	65.58 °C	Total nmol	40.42 nmol
Skala	50 nmol	MW	10995.20	Conc	1587.31ng/ul	Total ng	444445.69 ng
Purified	DSLIT	A260	49.3	OD	13.8	100 um stock	404.2

Table 2.6. Specifications of reverse *EtXTH1* primer

Oligo No	191113-1-7	GC	44%	Tm(basic)	70.94 °C	Total nmol	41.76 nmol
Skala	50 nmol	MW	11599.55	Conc	1730.05ng/ul	Total ng	484413.03 ng
Purified	DSLIT	A260	52.6	OD	14.7	100 um stock	417.6

2.4. DONOR AND ACCEPTOR SUBSTRATES

Presence and quantity of endotransglycosylase/hydrolase enzymes are determined by enzymatic activity assays therefore interactions of specific enzyme -substrates are analyzed by the usage of polysaccharide donors and oligosaccharides supplied from Megazyme International Ireland (Table 2.7. and Table 2.8.). Enzyme activity assay works by mimicking components of the plant cell wall using polysaccharides donors. XTH enzymes are associated with endotransglycosylation and endohydrolysis, therefore our purpose is to determine the specificity of the substrates and the role of the produced XTHs. *AtXTH30* (*Arabidopsis thaliana*) and *EtXTH1* (*Equisetum telmateia*) genes were supplied from Twist Bioscience, San Francisco, USA.

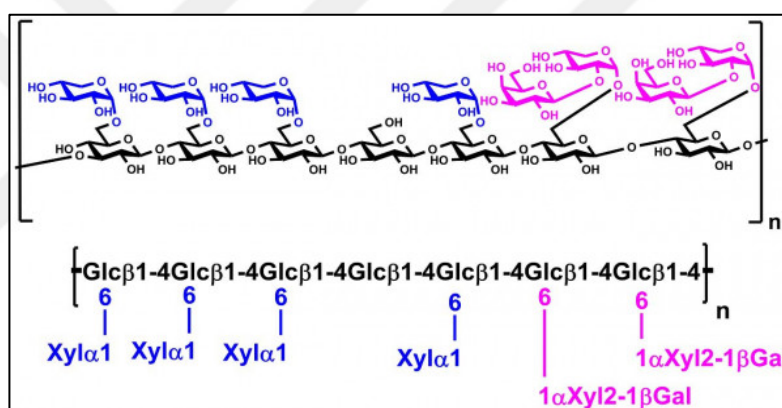


Figure 2.1. Representation of Tamarind Seed Xyloglucan (TXG) polysaccharide as donor substrate for enzymatic activity assays [3].

Table 2.7. List of donor and acceptor substrate, abbreviations, and catalog numbers

Polysaccharide Donor	Abbreviation	Catalog number
Tamarind seed xyloglucan	TXG	P-XYGLN
Oligosaccharide Acceptor	Abbreviation	Catalog number
Xyloglucan oligosaccharide	XGO	O-XGHON

3. METHODS

3.1. TARGET ENZYME SELECTION

Various codon-optimized genes *EtXTH1* from agriculturally and economically significant *Equisetum telmateia* and *AtXTH30* from the model plant *Arabidopsis thaliana* were predicted to have hydrolytic activity with xyloglucan endotransglycosylase activity, therefore selected genes were cloned into expression vector pPICZ α -C, then continue with heterologous expression in methylotrophic *Pichia pastoris*. Amino acid sequences of targeted plants for purification and characterization were retrieved from the NCBI database.

3.2. GLYCEROL STOCK PREPARATION

Preparation of glycerol stocks gives an opportunity to stabilize the bacteria and inhibit cell membrane damage, therefore for long-term storage, DH5 α *Escherichia coli* cells and *Pichia pastoris* cells were grown overnight to prepare glycerol stocks. 180 μ l 87 percent v/v glycerol and 820 μ l cell cultures were added to separate cryotubes up to a total volume of 1000 μ l., samples were inverted a few times and kept at -80 °C low-temperature freezer. For competent bacterial glycerol stocks, DH5 α *E.coli* cells grown overnight in 10 mL of low salt pH 7.5 Luria-Bertani (LB) medium at 37 °C with 150 rpm, and the overnight cultured 10 mL of *E.coli* cells were incubated in 100 mL LB medium and cultured to OD₆₀₀ value 0.4. by measuring every 2 hours, depending on the inoculum. Then, cells were centrifuged at 4°C 3000 x g for 10 minutes, and pellets were double washed with ice-cold 75 mM calcium chloride (CaCl₂) and 15 percent v/v glycerol. Then, aliquoted and stored at -80 °C.

3.3. EXPRESSION PLASMID PREPARATION

Amplification of the target gene *EtXTH1* was carried out by PCR using PCR Master Mix (ThermoFisher), template of expression target supplied from Twist Bioscience as plasmid powder, and reverse-forward primers. Size of the PCR products was checked by 1 percent agarose gel electrophoresis. Then, gel was extracted and purified of this DNA fragment

with 567 bp long. Plasmid powder of pTwist/*AtXTH30* resuspended in 20 μ l of nuclease-free water in separate microcentrifuge tubes. 2 μ l of resuspended plasmid transformed into 50 μ l of competent *E.coli* DH5 α cells to increase plasmid yield. Expression vector pPICZ α (reading frame C) transformed into DH5 α *E.coli* cells which will be used for heterologous expression in *Pichia pastoris*.

The next step consists of splicing the digested plasmids by ligation reaction, therefore entire plasmids contain ClaI restriction site at 5' end and XbaI restriction site at 3' end for proper cloning. Transformed DH5 α cells with pTwist/*AtXTH30* and pTwist/*EtXTH1* plasmids were grown in low salt LB agar plates with the final concentration of 100 μ g/ μ l of Ampicillin, and DH5 α cells that contain pPICZ α -C plasmids were grown in low salt LB agar plates with the final concentration of 25 μ g / μ l of Zeocin overnight. Selected single colonies from the agar plates were inoculated in 10 mL of LB Medium for 12-16 hours and isolated with Qiagen Plasmid Extraction Kit. Plasmid DNA concentration quantified by NanoDrop before stored at -20°C.

Plasmids consist of expression targets *AtXTH30* and *EtXTH1* and the expression vector pPICZ α -C were double digested by ClaI and XbaI restriction enzymes to cut plasmids at stated specific restriction sites for 1 hour incubation time that depends on the amount of DNA digested at 37°, and the restriction endonucleases inactivated at 65°C for 20 minutes. Double digested plasmids were run on 1 percent agarose gel at lower voltages (80-90V) for 1 hour to increase the resolution of bands at specific length in base pairs (*AtXTH30* 960 bp, *EtXTH1* 567 bp, digested pPICZ α -C 3296 bp). Bands were extracted with Macherey-Nagel Gel Clean-up Kit and extracted plasmid DNA fragments of *AtXTH30*, *EtXTH1* and pPICZ α -C purified and ligated using T4 DNA Ligase with a 2:1 molar ratio of insert:vector correlated with the concentration and size of the DNA samples and maximum reaction volume. Ligation reaction of sticky-end fragments incubated at 4°C overnight, then transformed into competent *E.coli* cells and grew the transformants containing pPICZ α -C/*AtXTH30* and pPICZ α -C/*EtXTH1* in LB agar plates with Zeocin.

3.3 PLASMID ISOLATION

Recombinant plasmids pPICZ α -C/*EtXTH1* and pPICZ α -C/*AtXTH30* were isolated from *E.coli* cells by plasmid isolation. Transformant DH5 α cells were grown in a 10 ml LB medium (pH 7.5) with the final concentration of 25 μ g/ μ l of Zeocin by following the instructions from the Invitrogen EasySelect *Pichia* Expression Kit Manual. Grown culture of transformants isolated with Qiagen Plasmid Isolation Midi Kit while cleared lysate and precipitation of isopropanol accomplished by centrifugation. Plasmid concentrations quantified by NanoDrop Spectrophotometer under OD260/280 ratio that indicates the purity of DNA.

3.4. BACTERIAL TRANSFORMATION

Expression vectors, expression targets, and ligation reactions were transformed into DH5 α cells in LB agar plates to stock for future use and blue-white screening.

Chemically competent *E.coli* DH5 α cells are used for heat shock transformation to replicate the plasmid into the cells. Competent DH5 α cells were thawed and 50 μ l of competent cells were mixed with 2 μ l of ligated plasmids, incubated on ice for 20 minutes. Cells were placed at 42°C for 45 seconds and replaced on the ice for two minutes. 950 μ l of SOC medium was added to tubes and incubated at 37°C and 150 rpm for 90 minutes. Then, the selection of transformed cells was achieved by plating the cells on LB agar plates with zeocin by spreading, incubating at 37°C for 16-18 hours. Additionally, PCR products of *EtXTH1* were cloned into pre-linearized Promega pGEM-T Easy vectors on LB/ampicillin/IPTG/X-Gal plates to identify the recombinant colonies by blue-white screening.

3.5. PHENOL EXTRACTION AND ETHANOL PRECIPITATION OF DNA

DNA samples re purified to lysis cells and break down redundant cell materials that would cause contamination. DNA samples were mixed with equal volumes of phenol and shaken for 15 seconds. Centrifugation was performed for five minutes at 2,000 rpm at room temperature to separate phases. Aqueous DNA phase as the top layer was extracted into a

new 1.5 ml tube and equal amounts of phenol: chloroform 5:1 was added and mixed. Then, centrifugation at 2,000 rpm for 5 minutes at room temperature to obtain separated DNA from the organic base. To test the success of phenol extraction and to concentrate the DNA sample, sodium acetate ethanol precipitation was performed. DNA samples were mixed with 0.1 volume of 3 M sodium acetate (CH_3COONa), 2.5 volumes of ice-cold absolute EtOH, and incubated overnight at - 20 °C. After the incubation, centrifugation was done at 4°C 14,000 rpm for 15 minutes. Cell pellets were washed with 2 volumes of 80 percent (v/v) EtOH, incubated for five minutes at RT (25 °C), and centrifugation at 4°C 14,000 rpm for 5 minutes. Excess ethanol was discarded by air-drying the DNA samples and resuspended in 20 µl of ddH₂O.

3.6. COMPETENT *PICHTIA PASTORIS* PREPARATION

Pichia pastoris cells of SMD1188H strain were grown in Yeast Extract Peptone Dextrose Sorbitol (YPDS) agar plates at 30°C for 2-3 days. When colonies were cultivated, a single colony was selected and grew overnight in YPD Broth at 30°C with 170 rpm shaking until optical density value OD₆₀₀ was reached at 0.6 - 1.0. Cells were centrifuged at 500 x g for five minutes at room temperature, and obtained cell pellets were resuspended in 9 mL of 10 mM bicine-NaOH, pH 8.3, three percent (v/v) ethylene glycol, five percent (v/v) (dimethyl sulfoxide) DMSO, 1 M sorbitol (BEDS solution) enhanced with 1 M DTT. Cells were incubated in a shaking incubator for 5 minutes at 30°C. Treated cells were centrifuged at 500 x g for 5 minutes and resuspended the pellets with 2 mL of BEDS solution. Aliquoted competent cells are ready for immediate use.

3.8. *PICHTIA PASTORIS* TRANSFORMATION AND SELECTION FOR POSITIVE COLONIES

5 µg of pPICZα-C/*EtXTH1* and 5 µg of pPICZα-C/*AtXTH30* linearized with SacI restriction endonuclease for 15 minutes at 37°C, heat-inactivated at 65°C for 5 minutes. Digested plasmids were detected by 1 percent agarose gel electrophoresis, extracted and purified by phenol to discard proteins, and the plasmids were precipitated by ethanol and resuspended in nuclease-free water, following the instructions for Phenol/Chloroform Extraction and Ethanol Precipitation. 80 µL of methylotrophic *P. pastoris* cells were

transferred into electroporation cuvettes with 5 µg of digested plasmids, and incubated on ice for 5-10 minutes. The electric pulse of 1.5 kV, 200 ohm (Ω), and 25 µF was applied by GenePulser Electroporation System and immediately added 1 mL of 1 M sorbitol without shaking at 30°C for 1 hour, and adding 1 mL of YPD medium to the tube and incubate for additional 1 hour by following the instructions on EasySelect *Pichia* Expression Kit Manual. Grown cells spread on YPDS agar plates with 100 µg/mL zeocin and incubated for at least 3 days at 30°C until colonies to appear. After 3 to 5 days, 5-10 colonies were selected and streaked on fresh YPDS agar plates containing 100 µg/mL zeocin for antibiotic selection of the positive colonies. Integration of plant genes in the *Pichia* genome was examined by Colony PCR.

A single colony from YPDS agar plate inoculated in 10 mL of Buffered Glycerol Complex (BMGY) medium at 30°C with 180 rpm for 12-16 hours until the cell density OD₆₀₀ was reached at 0.2-0.6. When reached, cells were pelleted by centrifugation at 3220 x g for 10 minutes, and pellets were resuspended in 10 mL of Buffered Methanol-Complex (BMMY) medium and checked until OD₆₀₀ is 1.0. *P. pastoris* cells were incubated at 22°C with 130-150 rpm shaking with induction of 1 percent v/v methanol per 24 hours for 3 to 5 days. For every sample, 100 µL of BMMY cultures were collected to measure protein concentration by separating them with SDS-PAGE Electrophoresis. End of the methanol induction for at least 3 days, *P. pastoris* cells were pelleted by centrifugation at 4°C 3220 x g for 10 minutes. Supernatants were collected to screen the presence of 6x-His-tagged protein by SDS-PAGE, Western blotting, and activity analysis for XTH enzymes.

3.9. COLONY PCR

Overnight grown 10-20 *P. pastoris* colonies were picked from LB agar plates for screening the transformants. 5 µL of 5 U/µL lyticase enzyme was added to 10 µL of cell culture, and incubated for 30 minutes at 30°C and incubated at -80°C for 15 minutes. Therefore, cells were able to be used as templates for PCR because their cell wall degraded. For colony PCR, 10 µL of PCR master mix, 1 µM of α -factor 5' and 1 µM of AOX1 3' primers, 2,2 µL of cell lysate, and 5 µL of nuclease-free water were used as PCR solution. PCR cycling parameters are set at 95°C for 5 minutes, denaturation at 95°C for 25 seconds, annealing at 64°C for 20 seconds, extending at 72°C for 60 seconds (40 cycles), and a final extension at

72°C for 5 minutes. Lastly, 1 percent agarose gel electrophoresis was run to screen the fragments by UV Transilluminator.

3.10. SDS-PAGE AND WESTERN BLOTTING

Induction samples gathered from small-scale production were screened by SDS-PAGE to obtain analytical separation of proteins [55]. In the first place, 900 μ L of centrifuged proteins in supernatants was mixed with 100 μ L of 10 percent TCA solution, mixed, and incubated on ice for one hour. Proteins were pelleted by centrifugation at 4°C 12000 x g for 10 minutes, the supernatant was discarded and the pellet was resuspended in 800 μ L of pure acetone with an overnight incubation at -20°C. Proteins were centrifuged at 4°C 6500 x g for 10 minutes to concentrate proteins as pellets and 500 μ L of -20°C pure acetone was added to the pellets twice, vortex, and centrifugation.

After the washing step, the samples were incubated on ice while the remaining acetone was discarded. Protein pellets were dissolved in 30 μ L of nuclease-free water and mixed with 1:1 2X Laemmli buffer, were dried by placing at 95°C for 7 minutes for heat denaturation and gel electrophoresis [56]. Supernatants of BMMY cultures were carried out with polyacrylamide gel electrophoresis that consists of 5 percent stacking and 12 percent resolving gel. The protein ladder and 20 μ L of protein samples were loaded on the gel and run at 180 V for 5-7 minutes, for samples to migrate into resolving gel, and voltage increased to 200 V and ran for approximately 35 minutes.

End of the SDS-PAGE, gels were stained with Coomassie blue stain and incubated on a rocking platform for 30 minutes, then destained with HPLC-grade water. SDS-PAGE gels are visualized by using ChemiDoc Imaging System and Image Lab software. To elute proteins, wet transfer was achieved by sandwiching the cassette with a sponge/filter paper/gel/membrane/filter paper/ sponge order. Membranes were blocked with a blocking buffer overnight at 4°C. After blocking, the membrane was incubated with 6x-His tag antibody dilutions. The membrane was washed with TBS-T for 45 minutes, discarded, and washed with TBS-T additional two times for 15 minutes each time. Membrane was washed with TBS for 10 minutes and to detect the chemiluminescence signals, screening solutions were mixed and poured onto the membrane. Finally, membranes are visualized by using ChemiDoc Imaging System and Image Lab software.

3.11. HETEROLOGOUS PROTEIN PRODUCTION

A single positive colony was picked for heterologous expression after positive colonies were grown on YPDS agar plates with 100µg/ml of zeocin at 30°C for 3 days. Inoculated a transformant in 10 mL of BMGY medium at 30 °C 180 rpm for approximately 20 hours. Then, the cell culture is transferred into 250-750 mL of BMGY medium and grown at 30°C for 20-22 hours at 180 rpm. The optical density of cells was measured until reaching 4.0-6.0. When OD₆₀₀ reached 4, BMGY culture was centrifuged and transferred to BMMY medium with an OD₆₀₀ of 1.0. Cells were incubated at 22°C 135 rpm for 3-5 days. For every 24 hours, methanol induction was achieved by adding 1 percent (v/v) methanol to metabolize by *P. Pastoris*.

End of the methanol induction, cell cultures were pelleted by centrifugation at 4°C 3220 x g for 10 minutes. Collected supernatants used for SDS-PAGE, Western blotting, and activity analysis for XTH enzymes.

3.12. PREPARATION OF DONOR AND ACCEPTOR MOLECULES

Acceptor substrate, xyloglucan oligosaccharides (XGOs), were fluorescently labeled with fluorescent sulforhodamine B, and excess dye was discarded with acetone. Labeled acceptors were dried and resuspended in 20 percent (v/v) EtOH. Stock solutions are stored at -20°C and diluted substrates are kept at 4°C until use. Concentrations were measured by NanoDrop at 566 nm. Then, stock solutions were diluted with ultrapure water for proper screening and studying of substrate specificities. Donor substrate, tamarind seed xyloglucan (TXG), was prepared with 0.8 percent (w/v) concentration. Polysaccharides were diluted with ammonium acetate buffer (pH 6.0) which yielded 0.4 percent donor concentration.

3.13. ENZYME CHARACTERIZATION AND ACTIVITY ANALYSIS

In this thesis, the endotransglucosylase activity with hydrolase activity of XTH enzymes was predicted to detect and measure by an enzyme activity test. Donor polysaccharides and

labeled acceptor oligosaccharides were used to imitating the cell wall of the plant. In PCR tubes, 10 μL of donor substrate, 1 μL of acceptor substrate, and 2 μL of enzyme solution that are heterologously expressed were added separately. Then, completely mixed before incubating at 30°C for a time. Incubation time for colony screening responses and donor-acceptor couple reactions was 24 hours. This investigation aimed was to carry out substrate specificity calculations and pH optimization reactions based on the enzyme's reaction potential and concentration. Depending on the donors chosen, enzyme kinetic reactions were incubated for varying amounts of time. For essential donor-acceptor couples, the number of repeats was extended beyond 3 until the R2 value for the slope of linear equations was greater than 0.95. The activity of the enzyme is terminated at the end of the incubation of the reaction by adding 6 μL of formic acid to lower the pH value of the reaction. The reactions were then examined, and the amounts of generated hybrid products were identified using HPLC analysis.

3.14. HPLC ANALYSIS

An "Agilent 1100 Series" high-pressure liquid chromatography (HPLC) device, with an integrated FLD fluorescence detector, was used to measure enzyme activity. The sulforhodamine fluorescent dye employed in the detection step has a 570 nm excitation wavelength and a 590 nm emission wavelength. Separation of fluorescently labeled polysaccharides and oligosaccharides was performed using a BioSep-SEC 4000 column, 75x7,80 mm, with a mobile phase of 80 percent 0,1 M pH 6.0 ammonium acetate and 20 percent acetonitrile at a flow rate of 0.5 ml/min to analyze enzyme reactions. The ChemStation software tool was used to complete the evaluation. The total reaction volume was 30 μL , therefore 11 μL of distilled water was added to 19 μL of reaction. For measurements, 15 μL of the reaction was injected by the equipment.

3.15. *IN SILICO* SEQUENCE ANALYSIS

DNA sequences and transcriptome sequences were used in *In silico* analyses to investigate the XTH genes in detail at transcriptomic and protein levels. The 1.000 Plants Project, used in this research as a target database, acquires sequence for 1000 plants, including *Equisetidae* subclass members; *Equisetum hymale* and *Equisetum diffusum*. Transcriptome

data of *Equisetum telmateia* was gathered from previous studies by M. Seven [57]. By BLAST, the *Equisetidae* subclass was selected as the target organism to determine local similarities, depending on the expected value (e-value) to validate the result whether it's a random match or not. Chosen similar protein sequences (i.e. gnl|onekp|JVSZ_scaffold_2001367 indicates an XTH gene of *Equisetum hymale* with 2001367 GenBank number) downloaded into Sublime Text to sort, generated a BLAST database, then blasted by Blast2GO. The GenBank accession numbers of *Arabidopsis thaliana* XTH genes are given in Table 3.1. During the alignment, four-letter codes such as JVSZ, CAPN indicate the species of the plants from the 1,000 Plants Project; JVSZ indicates *E. hymale* and CAPN indicates *E. diffusum*.

Table 3.1. Accession numbers of selected *A. thaliana* XTH genes

Name of the gene	GenBank accession number
<i>AtXTH1</i>	AT4G13080
<i>AtXTH2</i>	AT4G13090
<i>AtXTH3</i>	AT3G25050
<i>AtXTH4</i>	AT2G06850
<i>AtXTH5</i>	AT5G13870
<i>AtXTH6</i>	AT5G65730
<i>AtXTH7</i>	AT4G37800
<i>AtXTH8</i>	AT1G11545
<i>AtXTH9</i>	AT4G03210
<i>AtXTH10</i>	AT2G14620
<i>AtXTH11</i>	AT3G48580
<i>AtXTH12</i>	AT5G57530
<i>AtXTH13</i>	AT5G57540
<i>AtXTH14</i>	AT4G25820
<i>AtXTH15</i>	AT4G14130
<i>AtXTH16</i>	AT3G23730
<i>AtXTH17</i>	AT1G65310
<i>AtXTH18</i>	AT4G30280

<i>AtXTH19</i>	AT4G30290
<i>AtXTH20</i>	AT5G48070
<i>AtXTH21</i>	AT2G18800
<i>AtXTH22</i>	AT5G57560
<i>AtXTH23</i>	AT4G25810
<i>AtXTH24</i>	AT4G30270
<i>AtXTH25</i>	AT5G57550
<i>AtXTH26</i>	AT4G28850
<i>AtXTH27</i>	AT2G01850
<i>AtXTH28</i>	AT1G14720
<i>AtXTH29</i>	AT4G18990
<i>AtXTH30</i>	AT1G32170
<i>AtXTH31</i>	AT3G44990
<i>AtXTH32</i>	AT2G36870
<i>AtXTH33</i>	AT1G10550

Selected *Arabidopsis thaliana* XTH genes are *AtXTH1*, *AtXTH2*, *AtXTH3*, *AtXTH4*, *AtXTH5*, *AtXTH6*, *AtXTH7*, *AtXTH8*, *AtXTH9*, *AtXTH10*, *AtXTH11*, *AtXTH12*, *AtXTH13*, *AtXTH14*, *AtXTH15*, *AtXTH16*, *AtXTH17*, *AtXTH18*, *AtXTH19*, *AtXTH20*, *AtXTH21*, *AtXTH22*, *AtXTH23*, *AtXTH24*, *AtXTH25*, *AtXTH26*, *AtXTH27*, *AtXTH28*, *AtXTH29*, *AtXTH30*, *AtXTH31*, *AtXTH32*, *AtXTH33*; *Oryza sativa* XTH genes are *OsXTH1*, *OsXTH2*, *OsXTH3*, *OsXTH4*, *OsXTH5*, *OsXTH6*, *OsXTH7*, *OsXTH8*, *OsXTH9*, *OsXTH10*, *OsXTH11*, *OsXTH12*, *OsXTH13*, *OsXTH14*, *OsXTH15*, *OsXTH16*, *OsXTH17*, *OsXTH18*, *OsXTH19*, *OsXTH20*, *OsXTH21*, *OsXTH22*, *OsXTH23*, *OsXTH24*, *OsXTH25*, *OsXTH26*, *OsXTH27*, *OsXTH28*, *OsXTH29*, *OsXTH30*, *OsXTH31*; *Equisetum* family XTH genes are (*Equisetum diffusum*, *Equisetum hymale*) CAPN_2041016, JVSZ_2128190, CAPN_2039510, CAPN_2040035, CAPN_2039437, CAPN_2010178, CAPN_2004961, JVSZ_2011101, CAPN_2010498, JVSZ_2128529, JVSZ_2001367, CAPN_2001217, JVSZ_2008004, JVSZ_2012360, JVSZ_2009822, JVSZ_2012713, CAPN_2039395, CAPN_2039763, JVSZ_2130542, JVSZ_2007998, CAPN_2038789, JVSZ_2128837, CAPN_2009120, JVSZ_2012432, JVSZ_2131508, EfHTG1; and *Equisetum telmateia* XTH genes from previous studies [57] were used in *in silico* analyses.

3.15.1 Multiple Sequence Alignment and Phylogenetic Tree

Transcriptome data of selected plants were gathered from OneKP 1,000 Plants Project (<https://db.cngb.org/onekp/>). In FASTA format, DNA sequences of XTH genes from selected plants were blasted by Blast2GO Software and similar sequences were identified. To generate a Multiple Sequence Alignment (MSA), 33 XTH genes of *Arabidopsis thaliana*, 31 XTH genes of *Oryza sativa*, 67 XTH genes of multiple *Equisetum* species (23 from *E. hymale*, 22 from *E. diffusum*, and 22 from *E. telmateia*) were used by Clustal Omega. Based on the multiple sequence alignment, a phylogenetic tree was created likewise. A phylogenetic tree was visualized by Dendroscope 3.7.6 Program, a radial phylogram was obtained using sequence files in dnd format. The phylogenetic tree saved as an image (png file) to separate the plant XTH genes depending on their groups. (i.e. demonstrate XTH genes of *Arabidopsis thaliana* from *AtXTH12* to *AtXTH26*, belongs to Group II, colored with green in phylogenetic tree). *AtXTH30* genes were chosen from Group III-B and *EtXTH1* were chosen from the Ancestral Group.

3.15.2. Signal Peptide & N-linked Glycosylation Site Prediction & DREME

Signal peptides and cleavage sites of protein sequences that encode for XTH of *Arabidopsis thaliana* and *Equisetum* family members were predicted by SignalP Program. XTH sequences without signal peptides were separated due to lack of N-glycosylation machinery. N-linked glycosylation sites, which are mostly NXS/T sites, were detected by using NetNGlyc 1.0 Software Program. Short, ungapped, and highly conserved protein motifs were predicted by DREME (Discriminative Regular Expression Motif Elicitation) analysis in MEME Suite (<https://meme-suite.org>)

3.15.3. Swiss Modeling & Swiss PDB Viewer

Signal peptides were obtained from SignalP 6.0 software and discarded from the rest of the sequences, known as mature proteins used for protein modeling. By using mature protein sequences of *AtXTH30*, *AtXTH30* was modeled by using 1.UZ.1 A (XET in complex with the XLLG) as a template in SWISS-MODEL (<https://swissmodel.expasy.org>). To determine the location of highly conserved protein motifs on the modeled 3-D protein

structure, DREME analysis was carried out in MEME Suite. Due to DREME results, highly conserved regions were checked in amino acid sequences and conserved motifs were colored in the predicted protein structure by using the Swiss PDB Viewer.



4. RESULTS

4.1. TARGET ENZYME SELECTION

AtXTH30 and *EtXTH1* plasmids were supplied in pTwist vectors by Twist Bioscience. Due to having double XbaI restriction sites in *EtXTH1*, the second XbaI restriction site was removed and the plasmid was codon optimized. By using newly designed primers, PCR was carried out with *EtXTH1* plasmid. PCR was problematic depending on the primer lengths and annealing temperatures. The protein weight and activity of expressed recombinant enzymes from *AtXTH30* and *EtXTH1* are suggested estimating after optimization. Amplification of *EtXTH1* fragments by PCR gave bands with 567 base pairs long which were observed by 1 percent agarose gel electrophoresis. Based on the differences between the melting temperatures of the primers, Gradient PCR was applied with specific melting temperatures as 56.6°C, 56.7°C, 57.9°C, 59.8°C, 62.1°C, 64.0°C, 65.3°C, 66°C, and 68°C. Supplied primers were 36 base pairs in length, however, recommended length of PCR primers is around 18-30 nucleotides according to Thermo Fisher. Extraction of PCR products ligated into expression vectors, hereby 6X-His tag was attached to the 3' end of the *EtXTH1* gene and the signal peptide was attached to the 5' end. Codon optimized sequences of *AtXTH30* genes were suitable for cloning into expression vector pPICZα-C, and PCR fragments were 960 bp in length.

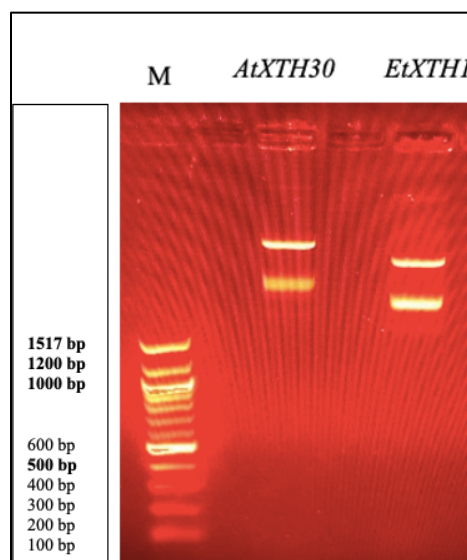


Figure 4.1. Uncut plasmid DNAs of *EtXTH1* and *AtXTH30*. M) 100 bp DNA Ladder.

To enhance the amount of plasmid, the received genes were firstly heat-shocked and transformed into DH5 α *E. coli* cells. *Pichia pastoris* transformation was achieved by electroporation, to assure the transformation, PCR was done with α -factor 5' and AOXI 3' primers.

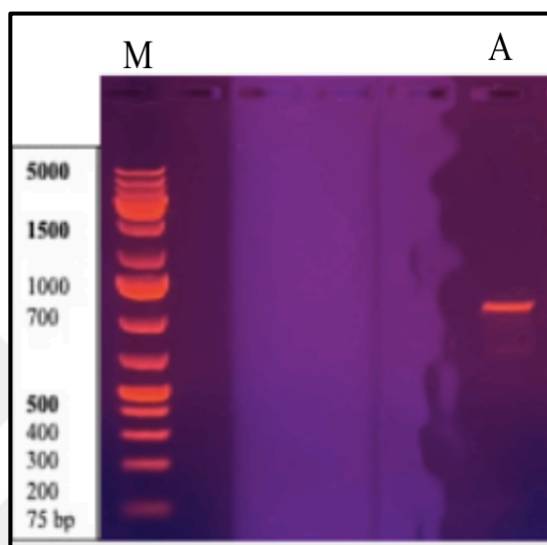


Figure 4.2. PCR products of *AtXTH30* on 1 percent agarose gel electrophoresis
M) GeneRuler 1kb plus DNA ladder A) *AtXTH30*

4.2. PROTEIN PRODUCTION AND CHARACTERIZATION

4.2.1. Expression Plasmid Preparation

Codon optimization of the plasmids with genes of interest and cloning of the sequences into pTwist vectors was performed by Twist Bioscience, USA. For the 5'- end of the sequence, the ClaI restriction recognition site was introduced and for the 3'- end of the sequence, the XbaI restriction recognition site was introduced. In *P. pastoris*, original sequence codons might be infrequent, and the yeast cell's deficiency of aminoacyl-tRNA might limit the synthesis of premature proteins. Due to slowdown, there could be issues with protein synthesis or folding. Due to codon optimization enhancing levels of expression, it will be preferable to utilize codon-optimized sequences of the XHT genes as heterologous expression techniques would be used to produce enzymes.

Plasmids which consist of target genes were firstly transformed into *E. coli DH5α* cells, grown and selected on LB agar plate with ampicillin (pTwist) or LB agar plate with zeocin (pPICZα -C). Bacteria containing recombinant plasmids were cultured overnight in LB + ampicillin or LB + zeocin liquid mediums. For the preparation of fifteen percent (v/v) glycerol stock, saturated cultures were used. Plasmid isolation was carried out with the rest of the overnight cultures. Genes of interest in pTwist vectors were cut from the plasmid by ClaI-XbaI restriction enzyme digestion.

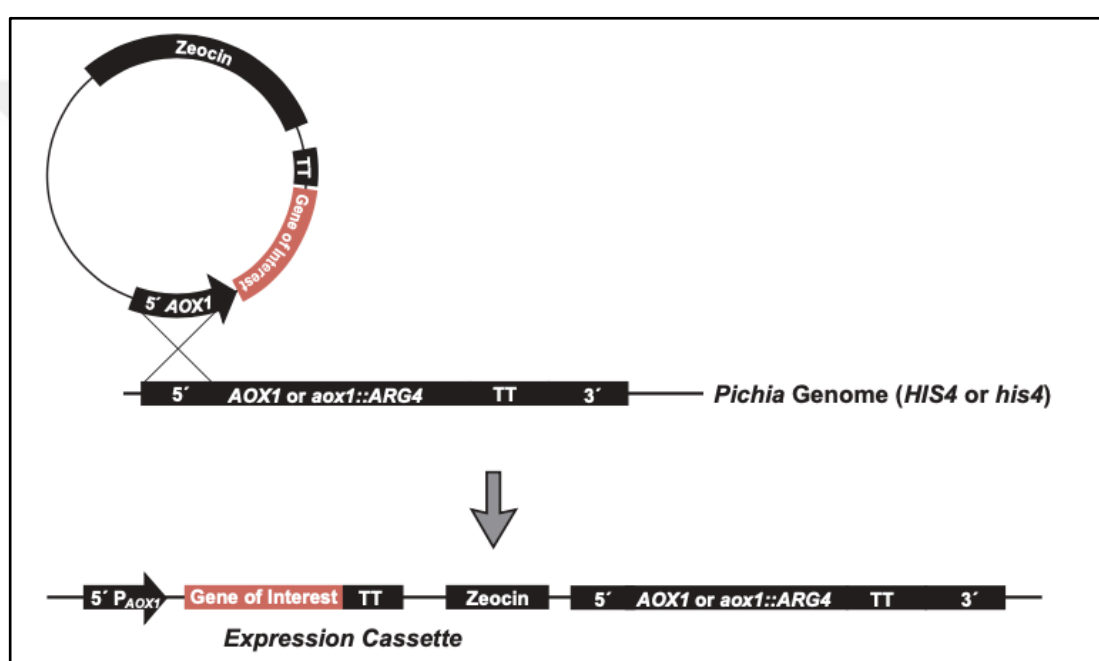


Figure 4.3. The expression cassette of *Pichia pastoris* genome [58]

Results of restriction enzyme digestion were separated on 1 percent w/v agarose gel by electrophoresis and *AtXTH30* and *EtXTH1* genes were extracted and purified from the gel using the Macharey-Nagel PCR clean-up gel extraction kit. The empty pPICZα-C plasmid was also digested using ClaI-XbaI restriction endonucleases and gel purified by Macharey-Nagel PCR clean-up Gel extraction kit. Digested genes of interest were ligated to digested empty plasmids using T4 DNA Ligase with a 2:1 vector: insert molar ratio based on the molarity calculations. Bacterial transformations were performed by mixing 2 µl of the ligation reaction with 50 µl of DH5α *E. coli* and several bacterial colonies that contain recombinant plasmids were selected from the LB agar plate (Figure 4.4.) and grown in LB

broth with zeocin overnight which plasmid isolation was performed by using these cultures.

All of the recombinant plasmids were linearized by their SacI restriction recognition site, precipitated by ethanol, and purified by phenol/chloroform extraction. The purified plasmid pellets were resuspended in nuclease-free water which was then used for *Pichia pastoris* transformation.



Figure 4.4. *AtXTH30*-pPICZ α -C *E.coli* transformant selection

4.2.2. Yeast Transformation

Similar to *Saccharomyces cerevisiae*, permanent transformants of *Pichia pastoris* can be generated via homologous recombination between the transforming DNA and homologous sections of the chromosome [59]. Even when expressed in numerous copies, such integrants exhibit exceptional stability in the lack of selection pressure. Depending on the host strain selected, it is simple to produce Mut⁺ or MutS recombinants via linearizing the recombinant plasmid by restriction digestion identified in the 5' AOX1 regions, in this case, SacI restriction endonucleases.

In pPICZ α -C heterologous yeast expression, the insert is separated from the cloning site by signal cleavage sites which are Kex2 and Ste13, and the c-myc epitope and polyhistidine tag are located after the cloning site. As an outcome, a protein produced in *P. pastoris*

using the pPicZ α -C expression system might have a polyhistidine tag and a c-myc epitope at the C-terminal as well as a factor signal at the N-terminus. The α -factor signal will primarily be removed from the Kex2 cleavage site, leaving Glu-Ala-Glu-Ala repetitions, after secretion of the produced protein outside of the cells. The α -factor signal will primarily be removed from the Kex2 cleavage site, leaving Glu-Ala-Glu-Ala repetitions, after secretion of the produced protein outside of the cells. Ste13 continues to cut Glu-Ala repeats afterward. The completely secreted protein may maintain a Glu-Ala-Glu-Ala repeat from Kex2 cleavage on its N-terminal end, and on its C-terminal end, it will have a c-myc epitope and polyhistidine tag.

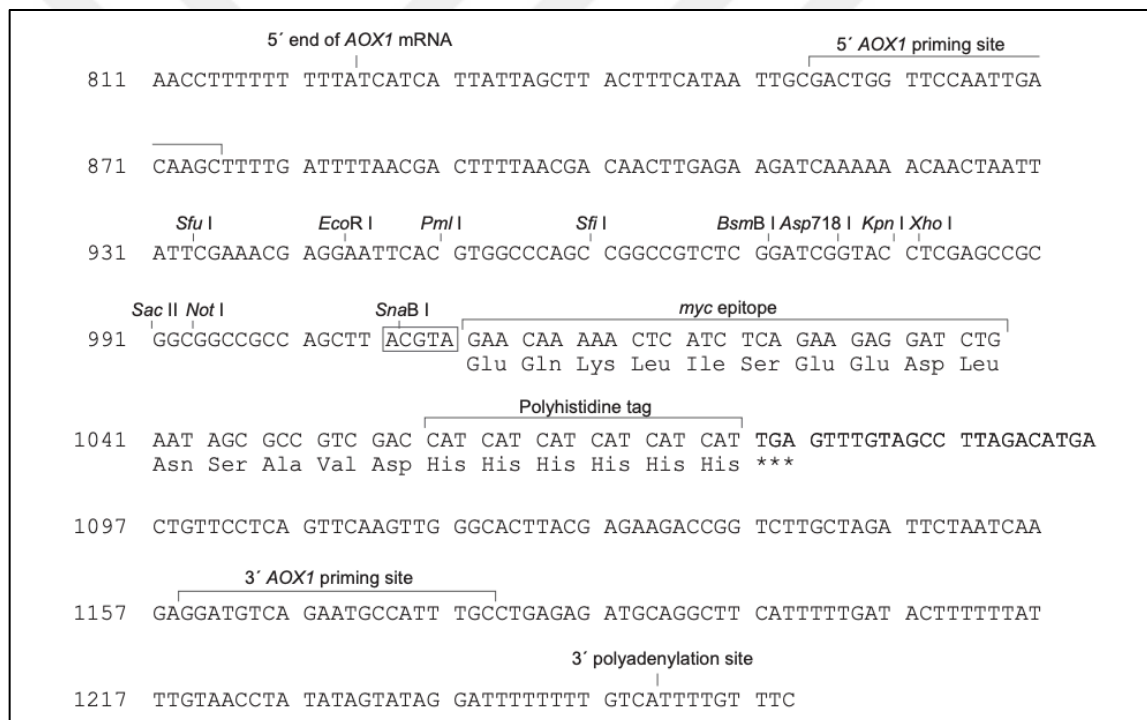


Figure 4.5. Sequence of the expression vector pPICZ α -C [60]

Several colonies with recombinant plasmids containing *AtXTH30* and *EtXTH1* were selected after transformation into the yeast genome. For *AtXTH30* gene, screening of transformants was carried out with 15 colonies. *Pichia pastoris* transformation by electroporation method did not give any transformants on LB agar plates with zeocin on the first try. Therefore, electroporation parameters; the voltage, the length, and the number of the pulses were changed. To prevent the arcing, linearized plasmids were resuspended in

nuclease-free water to decrease the salt concentration, rather than buffers. SacI linearized pPICZ α -C/*AtXTH30* plasmids electroporated into competent *Pichia pastoris* and grown on YPDS agar with zeocin. Empty expression vector pPicZ α -C was likewise linearized by the SacI restriction endonuclease and transformed into *P. pastoris* cells as a background control. In background control and non-transformant SMD1168H cells grow as the positive control, and no colonies were defined. Transformant pPICZ α -C/*AtXTH30* colonies were grown in YPDS agar with zeocin.

4.2.3. Positive Colony Selection and Screening

The expression of the gene after the AOXI promoter on the expression plasmid is caused by the induction of transformant yeast cells with methanol, which yields a protein product. It is not predicted to see any protein products expressed with a 6X-his tag when there is no insert after the promoter. It was intended to determine whether any background protein being synthesized and generated by *P. pastoris* which was 6X-his tagged and interacts with affinity purification was detected by screening different transformants. Additionally, it was expected to confirm the *P. pastoris* transformation procedure by running a colony PCR and examining the results for the completeness of the linearized plasmid in the chromosome inside the cells.

Transformant colonies were cultivated overnight at 30°C with 180 rpm shaking in 10 mL BMGY medium until the OD₆₀₀ value was 2–5. The 10 mL BMMY culture was then begun using cells with an OD₆₀₀ value of 1.0. Per 24 hours for five days, one percent (v/v) methanol was added to carry out gene expression. After day 5, centrifugation was used to separate the cells from the supernatant. Using the TCA and acetone precipitation procedure, proteins in sample supernatants were precipitated, concentrated, and subjected to SDS-PAGE and Western blotting analysis. Additionally, 5' AOXI and 3' AOXI primers were employed in colony PCR experiments using induced cells. For enzyme activity analysis, supernatants after 5 days of methanol induction were collected from the 50 mL Falcon tubes. 2 μ L of each supernatant was mixed with 10 μ L TXG and 1 μ L XGO-SRs incubated for 16 hours at 30°C to select the colonies that produce the most active enzyme. The reaction ended by adding formic acid. HPLC analysis was then completed on these enzyme reactions.

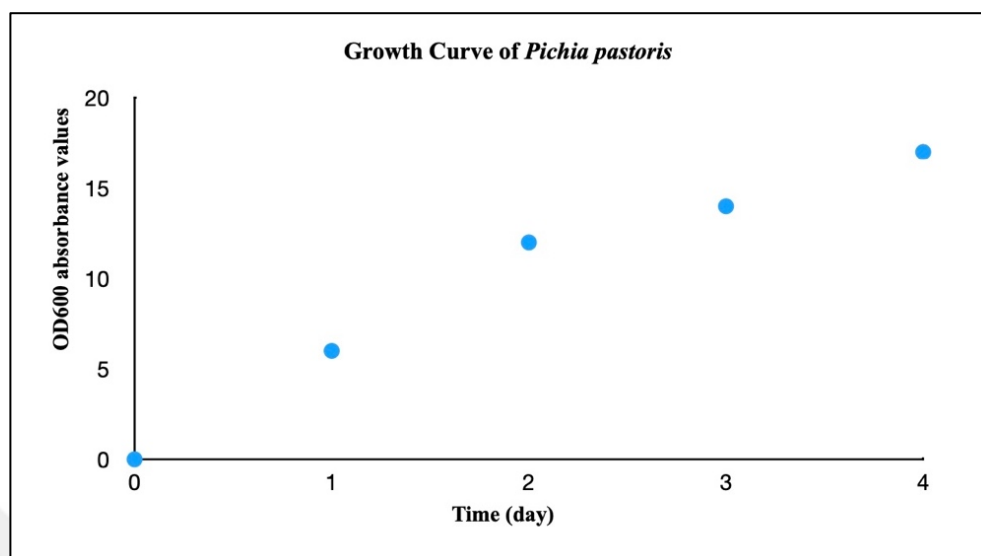


Figure 4.6. Growth curve of *AtXTH30-pPICZa-C* recombinant plasmids during methanol induction for four days.

In colony PCR results, a band-sized 960 bp for *AtXTH30* transformant colonies was expected. The gap between the 5' AOXI priming site and the 3' AOXI priming site, which is known to exist on vectors, is 593 bp. Colony PCR results show that homologous recombination was not effective because bands at 593 bp were not produced.

4.2.4. Detection of *AtXTH30* Protein

Produced enzymes could not be analyzed by ChemStation, because no peaks have been seen that show the presence of polysaccharides or oligosaccharides in the reactions. SDS-PAGE and Western blotting were used to verify the recombinant protein expression by positive colonies. SDS-PAGE gave indeterminate results based on obtained separation of protein mixtures. On the contrary, Western blotting gave negative results, which were absent of bands with expected size. In SDS-PAGE, thick bands were found to be somewhat larger than expected, which is probably due to glycosylation while expected protein bands were at 40 kDa.

Even though HPLC tests showed no activity, a few of the colonies still had bands of the identical sizes. Theoretically, treatment of membrane with Anti-6X-His antibody would

give protein bands that are equal to the expected, which provide the existence of the proteins. However, the results of our western blot do not imply the protein existence in culture for the *AtXTH30* enzyme.

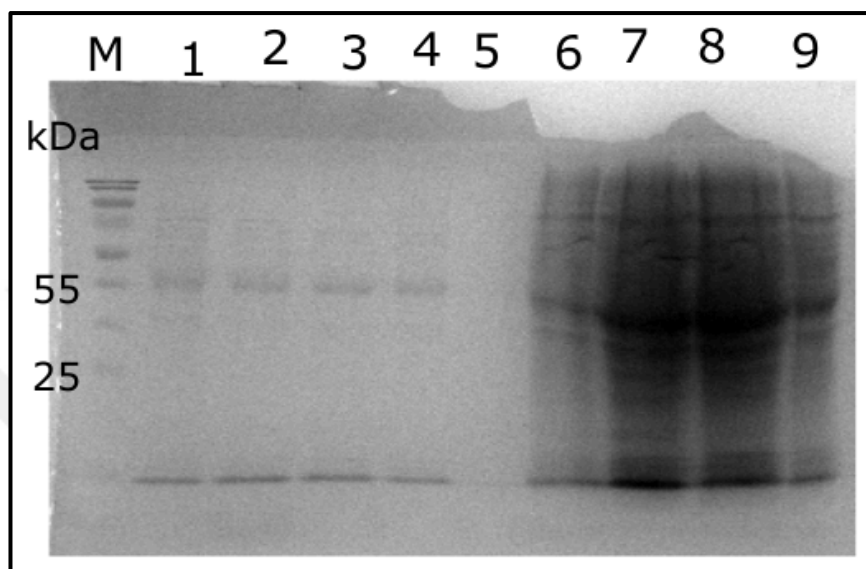


Figure 4.7. SDS-PAGE image of *AtXTH30* proteins from nine different colonies of small scale production run on twelve percent polyacrylamide gel.

M-PageRuler Protein Ladder

As previously confirmed, the active enzyme-producing colony was selected by checking on YPDS agar plates with zeocin and HPLC analyses. Protein bands exceeding 35 kDa were expected for enzymes, on the contrary, we were unable to detect the protein and purify it. In addition, normally in SDS-PAGE, 6X-His tag and c-myc epitope were attached to the structure of the enzyme because of using expression vectors for protein detection. However, most probably N-glycosylation is another reason for increasing the molecular weight and gives an over-thick band in screening. Undesirable structures were seen in the gel image, and the band structures were visible even if they were faint, also bands below 25 kDa admit contamination.

4.3. *IN SILICO* ANALYSIS OF XTH GENES

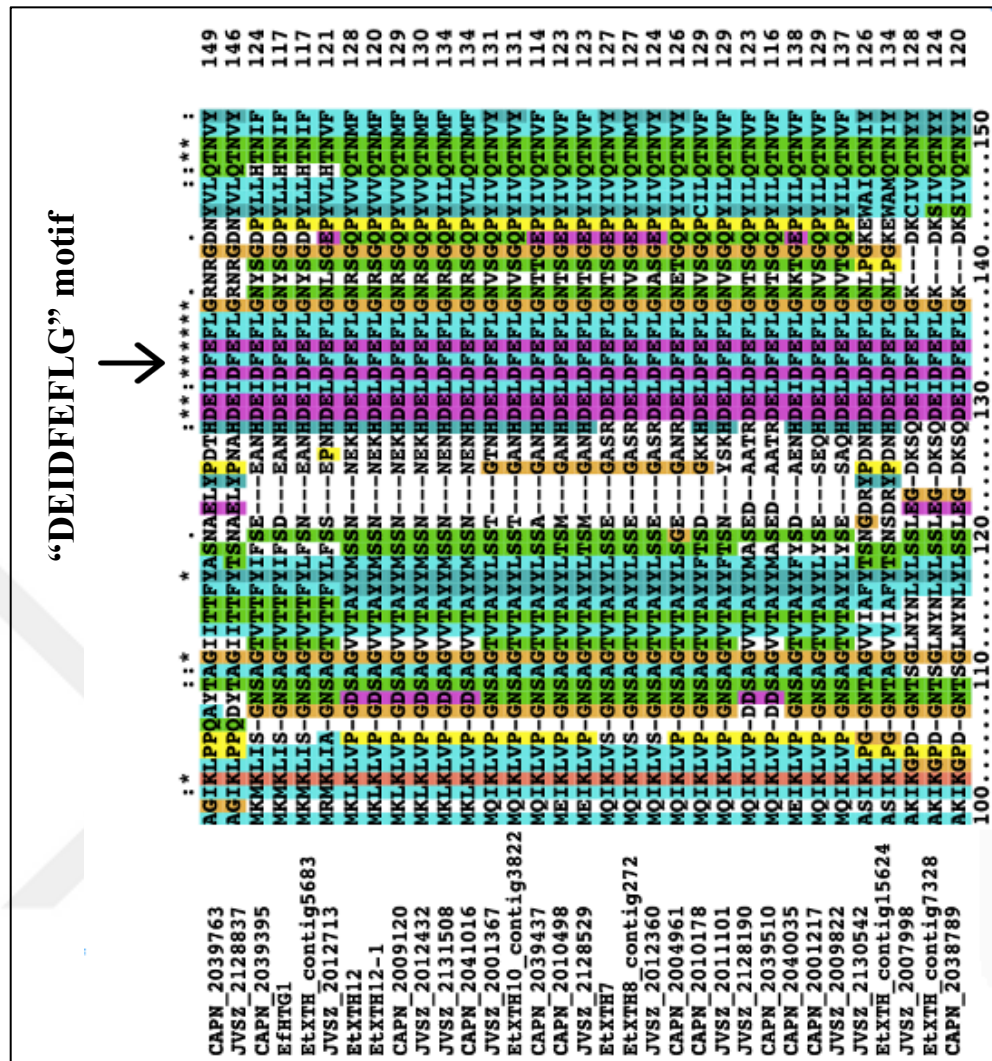
4.3.1. Multiple Sequence Alignment (MSA)

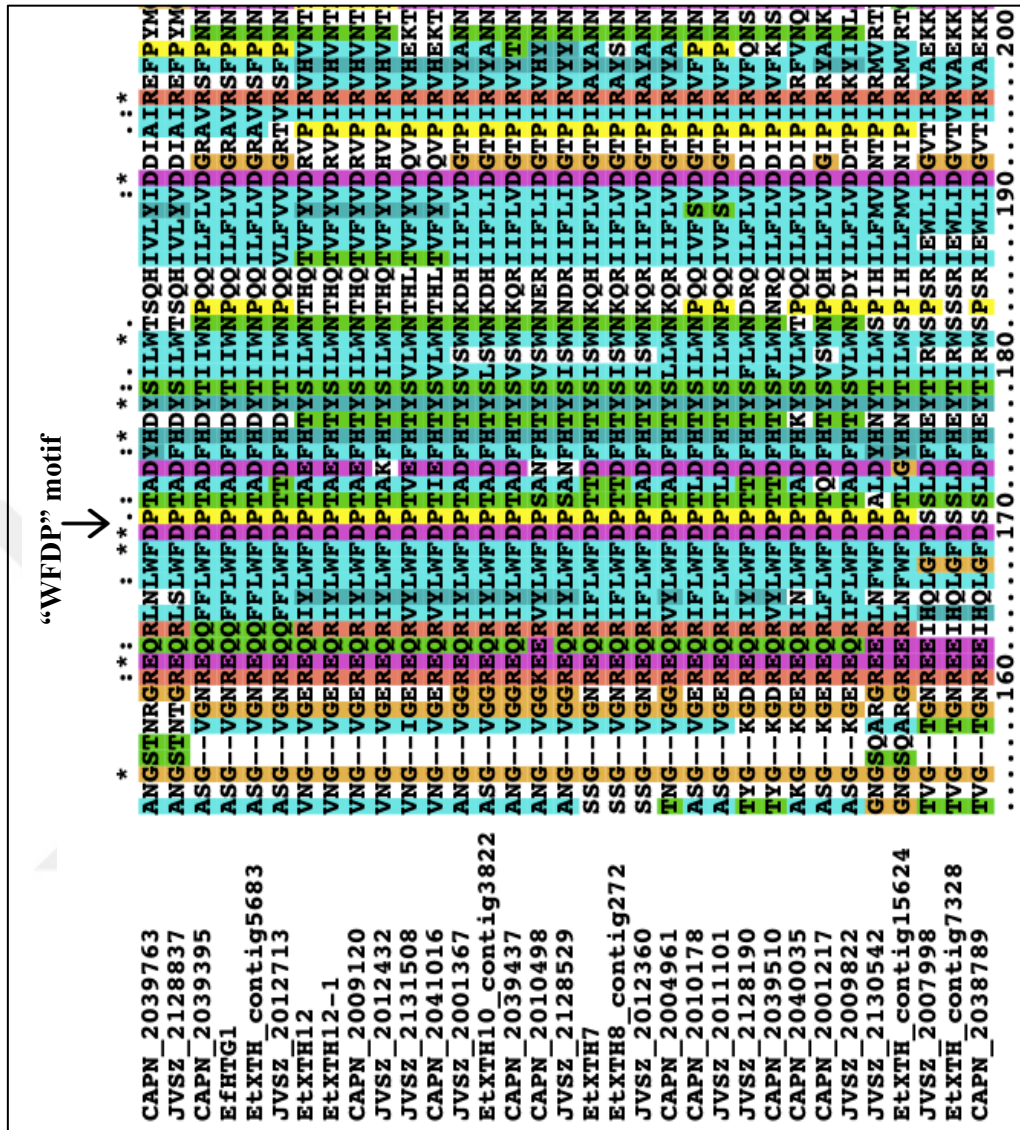
Simultaneously with laboratory experiments to secure the XTH enzyme of Equisetum, transcriptomes of three various Equisetum species; *E. hyemale*, *E. fluviatile*, and *E. diffusum*, were gathered as part of the One Thousand Plants Project (1KP). 33 XTH genes of *Arabidopsis thaliana* and *Oryza sativa* were gathered from the NCBI Nucleotide database (<https://www.ncbi.nlm.nih.gov/nucleotide>). Meanwhile, several studies about plant XTHs have been published in recent years. Publications were reviewed for certain plant XTH enzyme sequences, and the NCBI database (<https://www.ncbi.nlm.nih.gov/>) was used for scanning the database for sequences comparable to sequences of XTH from diverse plants.

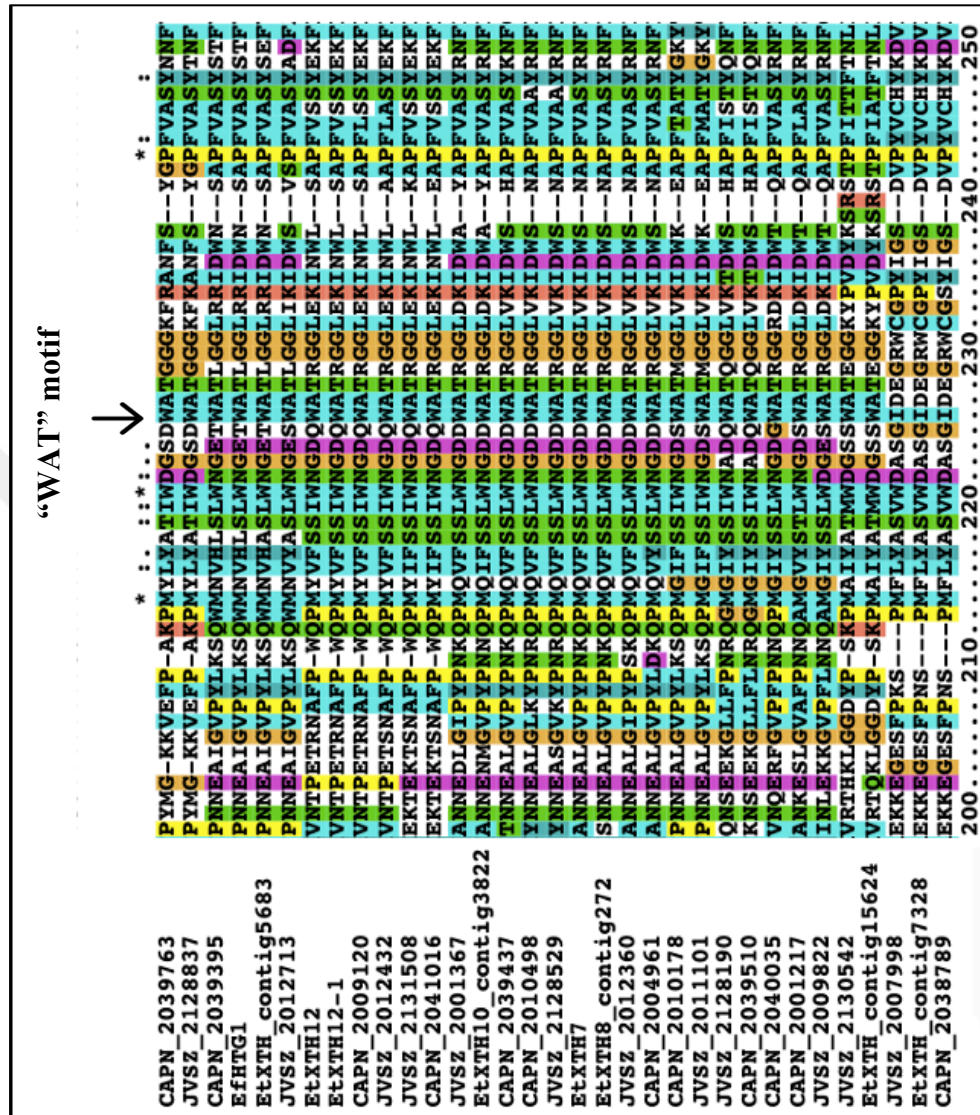
For Equisetum, the NCBI Protein Database has provided the amino acid sequences for 134 distinct genes, 70 genes from *Equisetidae* subclass. To encode every protein sequence, the FASTA format has been used. Because of their high degree of homology and common motifs in *Arabidopsis thaliana*, the multiple sequence alignment of AtXTHs sequences were shown in Figure 4.8, and the multiple sequence alignment of Equisetum XTH sequences visualized and were shown in Figure 4.9.

CAPN_2039763	MGKTI	CFVVLICLLVLVI	GLGGY	VEGK	CMLEPP	PPSTN	PLTEK	FPCTMY
JVSZ_2128837	MGKNV	CFVLVLTCCLV	---	GLGGY	VVGK	CMLEPP	PLT	PLTEKFPCTMY
CAPN_2039395	---	---	---	MKKKT	ASMLGLAF	GMLLIM	LASPK	LAIAG---
EfHTG1	---	---	---	MLGL	VFGMLVIM	LASPK	LAMAG---	---
EtXTH_contig5683	---	---	---	MLGL	AFGMLVIM	LASPK	LAMAG---	---
JVSZ_2012713	---	---	---	MKKKIG	---	MVLL	LGLFMII	IASPK-AEAN---
EtXTH12	---	---	---	MEF-V	ARIALS	---	VLVV-L	GLSALA-VVAQTVPGRSKF
EtXTH12-1	---	---	---	MEF-V	ARIALS	---	VLVV-L	GLSALA-VVAQTVPGRX---
CAPN_2009120	---	---	---	MEC-V	ARFVLS	---	VFVVM	GASSVA-AAAQTLPARSKF
JVSZ_2012432	---	---	---	MERRAA	TFLVF	---	FLGTAVA	AAVSLP-VAAQIIIPARSKF
JVSZ_2131508	---	---	---	MDSIL	SRSFAS	---	ILLIIAT	STTVWRYTEALPQIIIPGRTKF
CAPN_2041016	---	---	---	MESTMS	RLSAS	---	ILLIIAT	SAAVWCYTEALAOIIPGRPKF
JVSZ_2001367	---	---	---	MTSLPT	SPCLSL	ILIGL	GMALV	AYSSAKVNPVTISN---
EtXTH10_contig3822	---	---	---	MASLA	ASPCLSL	VIGL	GMVLV	ACSSAKVSPVATSS---
CAPN_2039437	---	---	---	MAALP	SP	---	MVAV	PLP-PHSS---
CAPN_2010498	---	---	---	MALLK	SFGC	---	LVALV	MASAVAVSIPMS---
JVSZ_2128529	---	---	---	MGLLK	SWSC	---	AVAFV	GMAFGVAMCIPSS---
EtXTH7	---	---	---	MAKYL	ASQCI	---	IALIL	VMASSSSSILKVAMAN---
EtXTH8_contig272	---	---	---	MATYL	APHCI	---	IALIL	VMASSSSSILKVAMAN---
JVSZ_2012360	---	---	---	MATYL	ATQC	---	IALIL	VMASS--SILPAAVAN---
CAPN_2004961	---	---	---	MVRYLA	ISAVGL	---	ASLFI	AMPSS--LAINSN---
CAPN_2010178	---	---	---	MEFCF	PLLHQR	---	VSQFL	LLSLLLLVVGNATAAPN---
JVSZ_2011101	---	---	---	MEFC	CPPLLHQG	---	VSQIF	LLSLLLLVVGNAAAASN---
JVSZ_2128190	---	---	---	MAALS	IF	---	LLLV	FSAACSSGPSLVDAN---
CAPN_2039510	---	---	---	MALGIC	---	---	FLLL	---
CAPN_2040035	---	---	---	MGDV	WFRS	GRRET	OYLL	LLFFIFVLWISSEVEAAGRASSF
CAPN_2001217	---	---	---	MASST	LN	LN	LVTS	---
JVSZ_2009822	---	---	---	MALP	SLN	RCRT	PT	THLSFFFSCFMAAAVVMVDG-LTTESFI
JVSZ_2130542	---	---	---	MIMN	MNR	---	NGR	LLILLIMLNCLITATA-RGVUS---
EtXTH_contig15624	---	---	---	MSI	IMN	MNT	TNMNM	NGRLLLIMLNCLITATA-RGVUS---
JVSZ_2007998	---	---	---	MSSEG	VA	VGGR	GGGEED	FLSSAAIPPSLLRTEKIR---
EtXTH_contig7328	---	---	---	MSSEG	VA	VG	---	---
CAPN_2038789	---	---	---	MVG	---	AGG	EEED	FLSSAAIPPSLLRTEKIR---
	1.....10.....20.....30.....40.....50							

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (a)

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (c)

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (d)

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (e)

CAPN_2039763	NNFVNGCPMTQDSK	---GNNVKPTGPCAN	---EFTQ	273
JVSZ_2128837	TNFIVNGCPMTQDSK	---GNNVKPVGPCAT	---EFTR	270
CAPN_2039395	STFVGDSCFDSAD	---SPCMASK	---WVNO	241
EfHTG1	STFVGDSCFDSAD	---SPCMASK	---WVNO	234
EtXTH_contig5683	SEFVGDSCFDSAD	---SPCMASK	---WVNO	234
JVSZ_2012713	ADFAADSCFDSAD	---SSCMATK	---WVNO	238
EtXTH12	EKFHADACGGG	---FCVPSSTT	---PWDO	244
EtXTH12-1	EKFHADACGGG	---FCVPSSTT	---PWDO	236
CAPN_2009120	EKFHADSCGGG	---SCAPSSST	---PWDO	245
JVSZ_2012432	EKFHADSCGGG	---SCAPSSST	---PWDO	246
JVSZ_2131508	EKFHANACQWSSATASSISTPPACASTGS	---TCPSSNST	---FWDO	262
CAPN_2041016	EKFHANACQWSSITASSGANTPPACASTGN	---TCPSSNST	---FWDO	262
JVSZ_2001367	RNFRANACVPSG	---TNPSNCLVNDAN	---NNN	252
EtXTH10_contig3822	RNFRANACVPSG	---GNPSNCLVHGAD	---NNN	252
CAPN_2039437	KNFQSNACIVPIG	---GTPNCLVHGAD	---NTN	235
CAPN_2010498	RNFKANACVVPKG	---GSASSCIANGSK	---NFN	244
JVSZ_2128529	RNFKANACVVPKG	---GSATSCIAYGAR	---NGN	244
EtXTH7	RNFRANACIVPKG	---GNPTTICIAQGAR	---NGN	248
EtXTH8_contig272	RNFRANACIVPKG	---GNPTTICIAQGAR	---NGN	248
JVSZ_2012360	RNFRANACVVPKG	---GNPTTICIAQGAR	---NGN	245
CAPN_2004961	RNFRANACVVPKG	---DNPNNCIAQGGV	---NAN	247
CAPN_2010178	GKYTELACES-KG	---EDTAASCAEGSW	---WNS	249
JVSZ_2011101	GKYTELACES-KG	---EDTSSCAEGAW	---WNS	249
JVSZ_2128190	QNFVIDACDLQNN	---RTSADPLCLSK	---WVNO	246
CAPN_2039510	QNFNIDTCNLQEN	---RTSADPLCLSK	---WVNO	239
CAPN_2040035	RNFNA-EACVVN	---RNMKKCKLASS	---SWDQ	258
CAPN_2001217	RNFSASTACVVN	---TNLNNCKQL	---WAHS	247
JVSZ_2009822	RNFQADACLVND	---ARDLRKQVFASSV	---TGQKKWSQ	266
JVSZ_2130542	TNLRLOGCPINPLLEDYIANKYNNNNNNVD	---ISSFQSNCGGYIFDEDEEDQ	---	275
EtXTH_contig15624	TNLRLOGCPINPLVEDYIADNNKNNNNIG	---ISNFQSCQGGYAFDEDEEDQ	---	283
JVSZ_2007998	KDVHIPSC	---	---	226
EtXTH_contig7328	KDVHIPSC	---	---	222
CAPN_2038789	KDVHIPSC	---	---	218
	250.....260.....270.....280.....290.....300			

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (f)

CAPN_2039763	IIPPSLSDTQKKSINWVKQNYMTNYCYEDTQRYPTLP	ECATQTKPSHTEAN
JVSZ_2128837	IIPPSLSDTQKQSINWVKQNYMTNYCYEDTQRYPTLP	ECATQTKPAHEEAN
CAPN_2039395	AAQSLSTSDASSIQVVRANLYKYDYCYDTKLYPNGFP	SECSNRGF
EtXTH1	AAQSLSTSDASSIQVVRANLYKYDYCYDTKLYPNGFP	RECSNRGF
EtXTH_contig5683	PAYQFLSTNDASSIQVVRANLYKYDYCYDTKLYPNGFP	AECSNRGF
JVSZ_2012713	PAYQFLSTNDASSIQVVRANLYKYDYCYDTKLYPNT-PP	IECQNRGF
EtXTH12	PSAWSLSPQEKQKYAVKSKLTVDYCYCTDKARYP-SLP	ADCMVHI
EtXTH12-1	PSAWSLSPQEKQKYAVKSKLTVDYCYCTDKARYP-SLP	ADCMVHI
CAPN_2009120	PSAWSLSPQEKRYAVKSKLTVDYCYCTDKARYP-SLP	ADCMVHI
JVSZ_2012432	PSAWSLSPQERRKYAVKSKLTVDYCYCTDKARYP-TTP	ADCKPSX
JVSZ_2131508	PRAWSLTPKEKLYAVKSKNTIYDYCTDVARYP-TTP	AECTLSF
CAPN_2041016	PRAWSLTPKEKSYVWVSKQTIYDYCTDVARYP-TTP	AECTLSF
JVSZ_2001367	AASKTMPKNFVSRVQRKFLIYDYCSDRARYP-SIP	TECTHNNL
EtXTH10_contig3822	VASKAIPKNYISRLRWQRKFLIYDYCSDRARYP-SIP	TECTHNNL
CAPN_2039437	VAITNTAQ-LYARLWVRQRKFLIYDYCSDHARYP-SIP	TECTHNTL
CAPN_2010498	TPIVDSQE-LNARLQWVRQRKFLIYDYCIDRARYP-STP	IECTKRKM
JVSZ_2128529	TPSIDSQE-LNARLQWVRQRKFLIYDYCIDRARYP-SIP	IECTRKKK
EtXTH7	SDVATNAQ-LHARLQWVRQRKFLIYDYCTDRARFS-TIP	VECAHSNL
EtXTH8_contig272	SDVTTNTQ-LNARLQWVRQRKFLIYDYCTDHARFS-TIP	IECTHNNL
JVSZ_2012360	SDVTTNTAQ-LNARLQWVRQRKFLIYDYCTDHARFS-TIP	IECTHNNL
CAPN_2004961	SNIADNGH-LNARLQWVRQRKFLIYDYCSDHARYN-TTP	IECAHNNL
CAPN_2010178	VNYQTMAEEDQQKLWVDNIIYDYCTDVARYP-EPP	VECVHQFATPLTSS
JVSZ_2011101	LNQQTMAEEDQQKLWVDNIIYDYCTDVARYP-EPP	VECVHQFATPLTSS
JVSZ_2128190	PYQQLTP-NQLANLQWVRQRKFLIYDYCKDQRFQGF	LECSINP
CAPN_2039510	PYQQLTP-NQLANLQWVRQRKFLIYDYCKDQRFQGF	LECSINP
CAPN_2040035	FTSSLNPQQAGSLHWVKKFMVYDYCTDLPRNP-TTP	TECTFN-LYIHIHIL
CAPN_2001217	NSNSX---QQS---LDWVKKNFIIYDYCTAQRP-TTP	PECAFN
JVSZ_2009822	QMPTS---NDKNKLEWVKNFMVYDYCTDSQRL-SPP	LECAFN
JVSZ_2130542	KGMWALKPQQVSSMRWARKNMYISYCDGKRYPLPL	ECSDNNSLVSRASR
EtXTH_contig15624	KGMWALRPQQVSSMRWARKNMYISYCDGKRYPPPL	ECSDNNSLVSRASR
JVSZ_2007998	-----	-----
EtXTH_contig7328	-----	-----
CAPN_2038789	-----	-----
	310.....320.....330.....340.....350.....360	

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (g)

CAPN_2039763	EANTKSEL	330
JVSZ_2128837	EANTKSEL	327
CAPN_2039395	-----	287
EFTG1	-----	280
EtXTH_contig5683	-----	280
JVSZ_2012713	-----	283
EtXTH12	-----	288
EtXTH12-1	-----	280
CAPN_2009120	-----	289
JVSZ_2012432	-----	290
JVSZ_2131508	-----	306
CAPN_2041016	-----	306
JVSZ_2001367	-----	297
EtXTH10_contig3822	-----	297
CAPN_2039437	-----	279
CAPN_2010498	-----	288
JVSZ_2128529	-----	288
EtXTH7	-----	292
EtXTH8_contig272	-----	292
JVSZ_2012360	LHVC	304
CAPN_2004961	-----	291
CAPN_2010178	TSSRASSPSSPSNTTIVTPNNSTSSISPSSASDGDNSSSPSSPFAINPYLATIAPAIATIFSSL	365
JVSZ_2011101	TSSASAPSSPSN-----PINSTSSISPSSATSGNDSSSPS-FAVHSYLATIAPS--IATIFSSL	357
JVSZ_2128190	-----	290
CAPN_2039510	-----	283
CAPN_2040035	HILMLVSRKIVIIILSYLPFFHSLPLEIROPIHFLLLVLSTHGEYA	354
CAPN_2001217	-----	285
JVSZ_2009822	-----	306
JVSZ_2130542	ASRKAKOHHRRHRRHNAKMRPLSSS-SQOEVVDEFP	359
EtXTH_contig15624	ASRKAVQHHRRHRRHNAKMRPLSSS-SQOEVVDEFP	368
JVSZ_2007998	-----	226
EtXTH_contig7328	-----	222
CAPN_2038789	-----	218
	360.....370.....380.....390.....400.....410.....420....	

Figure 4.9. Multiple Sequence Alignment of *Equisetum* XTH sequences (h)

4.3.2. Phylogenetic Tree of XTH Genes

Protein sequence alignments were deeply examined to build networks in phylogenetic approaches to detect and comprehend evolutionary relativities. By using the Clustal Omega software program, distinct phylogenetic trees were visualized and bootstrapped for illustration with the Dendroscope 3.7.6. program. Phylogenetic trees were modified for readability in radial phylogram form. The phylogenetic relevancy and differences among monocotyledon and dicotyledon plant species are demonstrated by an illustration of the phylogenetic trees. Moreover, bootstrap values from out of 1000 were collected and displayed on trees. A comparison between the selected monocots and dicot species could be seen in Figure 4.8. Six groups of XTH sequences—Group I, Group II, Group III-A, Group III-B, EG16, and Group IV were identified. Many distinct groups demonstrated a more comprehensive categorization.

33 XTH genes from *Arabidopsis thaliana* belonging to Group I, Group II, Group III-A, Group III-B, and Group IV. There are no AtXTH genes belonging to the EG16 group. *AtXTH30* and *EtXTH1* sequences, which are mRNA sequences, were obtained from the NCBI Nucleotide database (<https://www.ncbi.nlm.nih.gov/nucleotide>). 31 XTH genes of *Oryza sativa*, and 68 *Equisetum* XTH genes from multiple family members; 23 from *E. hymale*, 22 from *E. diffusum* and 22 from *E. telmateia* and 1 from *E. fluviatile* were used to generate the phylogenetic tree.

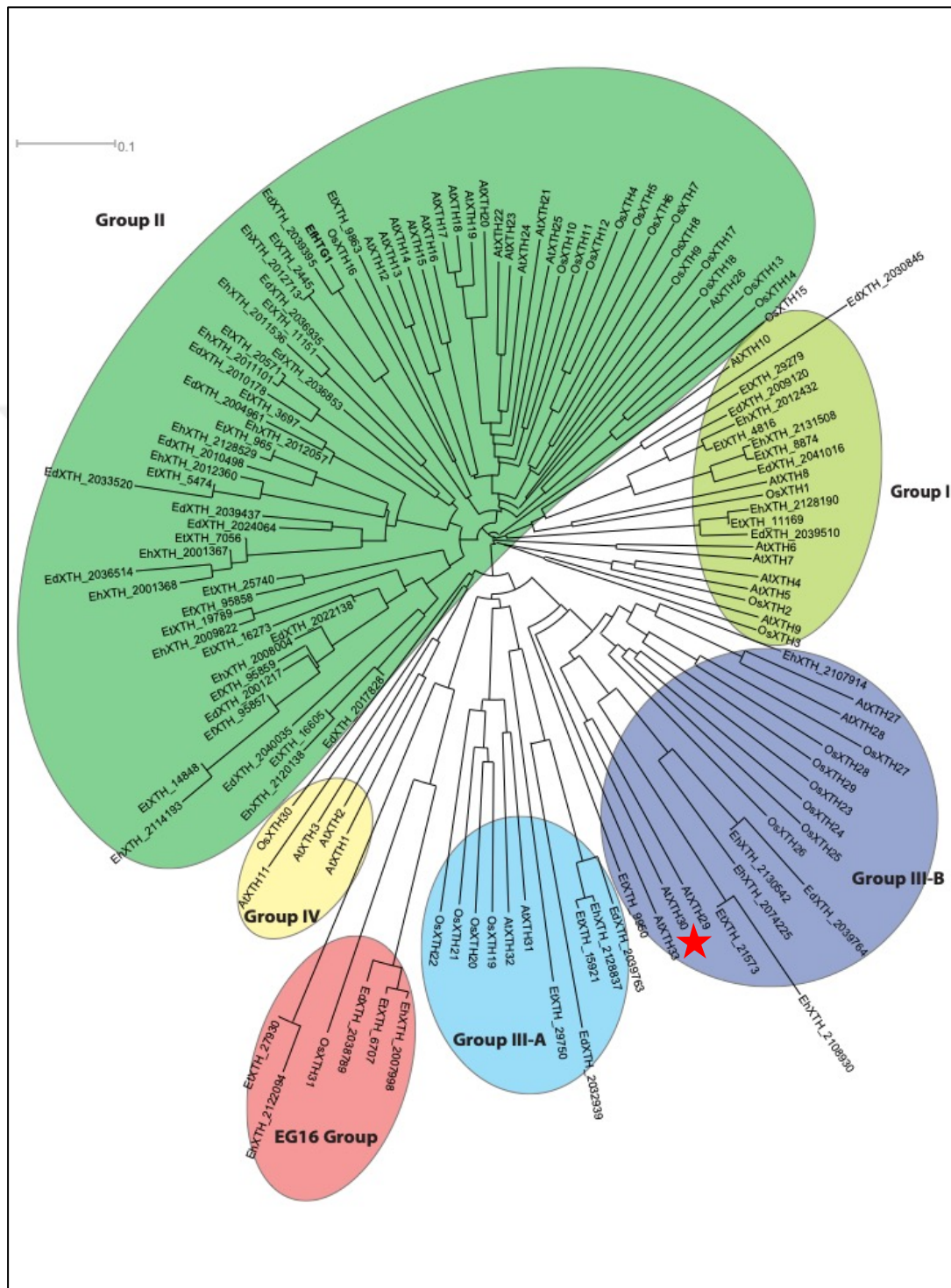


Figure 4.10. Phylogenetic tree of XTH genes of *A. thaliana*, *E. telmateia*, *E. diffusum*, *E. hymale*, *E. fluviatile* and *O. sativa*. Red star indicates *AtXTH30* gene used for molecular studies.

In recent studies [61], it is revealed that expression levels of XTHs could differ between the tissues of the plants; leaves, roots, floral, etc. Therefore, XTH expression levels and part of the plants that accumulate XTHs are important topics that should be investigated. Based on the gathered data from The Bio-Analytic Resource for Plant Biology, *AtXTH30* is frequently found in pollens and seeds, specifically in mature and young flowers (Figure 4.11).

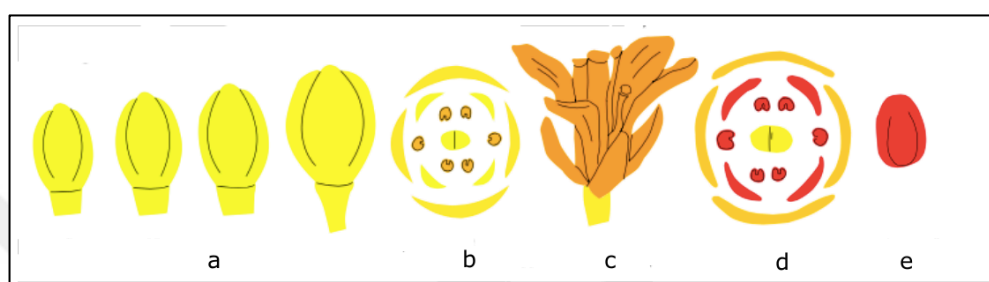


Figure 4.11. *AtXTH30* levels during flowering. a) During the flower stage, levels of XTH in begun to increase. From 28 to 75. b) At b stage, sep/pet/stam/carp parts of the plants, Tissue specific pollen germination *AtXTH30*.

Inflorescence indicated in letter a, the XTH level has a quite slow increment. Changes in floral organ lengths in *Arabidopsis thaliana* gave in a position-dependent manner across flower positions 1 to 10. In section a, flower positions were from 9 to 12. In position 12, flowers are shown just the newly opened in section b. [62] Depending on the phases of flower development in *A. thaliana*, describing the crucial points which characterize the start of a new phase and estimating its duration [63]. Based on Figure 4.11.b section b, the flower position 12 indicates the petals level with long stamens which is after the stigmatic papillae appear. The increment of *AtXTH30* level is increased in stamens, rather than sepals and petals. In Figure 4.11.c, at flower position 15, the stigma extends above long anthers, and at section d which is stage 15, stamens and petals painted with red indicate the high amounts of *AtXTH30* inside. And the last section e, demonstrates mature pollen which is the highest XTH level among their part.

This data got from the study [64] named “A gene expression map of *Arabidopsis thaliana* development” by Markus Schmid published in 2005. Examination of the complete seed in

the selected study, enables to detection of both genetic and epigenetic modulation inside the transcriptome of conserved RNA in both dry and imbibed seeds.

4.3.3. Signal Peptide and Mature Protein Determination

Arabidopsis thaliana, the higher plant with multigene families, has 33 genes for the XTH enzyme, different land plants, including vascular and non-vascular to monocots and dicots, have approximately 20 to 60 genes. The SignalP 6.1 Server online tool was used to identify the signal peptides and signal peptide cleavage sites of the protein sequences. During signal peptide prediction, Sec/SPI, Sec/SPII, Tat/SPI, Tt/SPII, and Sec/SPIII were examined. Sec/SPI known as standard signal peptides were collected and compared with each other. Specified signal peptides located beginning of the sequences were discarded, and the remaining protein-coding regions were identified. The identified protein-coding regions known as mature proteins, were used for modeling in SWISS-MODEL after the signal peptide tag was eliminated against the templates in the program. As a result, by SignalP, 20 out of 25 near-full length *Equisetum* XTH sequences contain a signal peptide (Sec/SPI), 33 out of 33 *Arabidopsis thaliana* XTH sequences contain signal peptide (Sec/SPI).

>AtXTH30 AT1G32170 xyloglucan:xyloglucosyl transferase, putative / xyloglucan endotransglycosylase, putative / endo-xyloglucan transferase, putative (XTR4);

MSKSSYNHIFILILCLCLRSSSAFTNLNTLSFEESLSPLFGDANLVRSPDDLVRLLLDY
TSGGFISSNMYQHGFYSSMIKLPADYTAGVVAFYTSNGDVFEKTHDELDIEFLGNIKGP
WRFQTNLYGNGSTHRGREERYRLWFDPSKEFHRYSLWTPHKIIFWVDDVPIREVIRNDAM
GADYPAKPMALYATIWDASDWATSGGKYKANYKFAPFVAEFKSFSLDGCSVDPIQEVPMDC
SDSVDFLESQDYSSINSHQRAAMRRFRQRFMYYSYCYDTLRYPEPLPECVIVPAEKDRFKE
TGRLKFGGTEARERRRNRQRRPEIEIESDPDDRKLL

Figure 4.12. The deduced amino acid sequence of *AtXTH30* contains a signal peptide and mature proteins. Green highlighted section indicated for signal peptide, cleavage site located between the position 23 and 24.

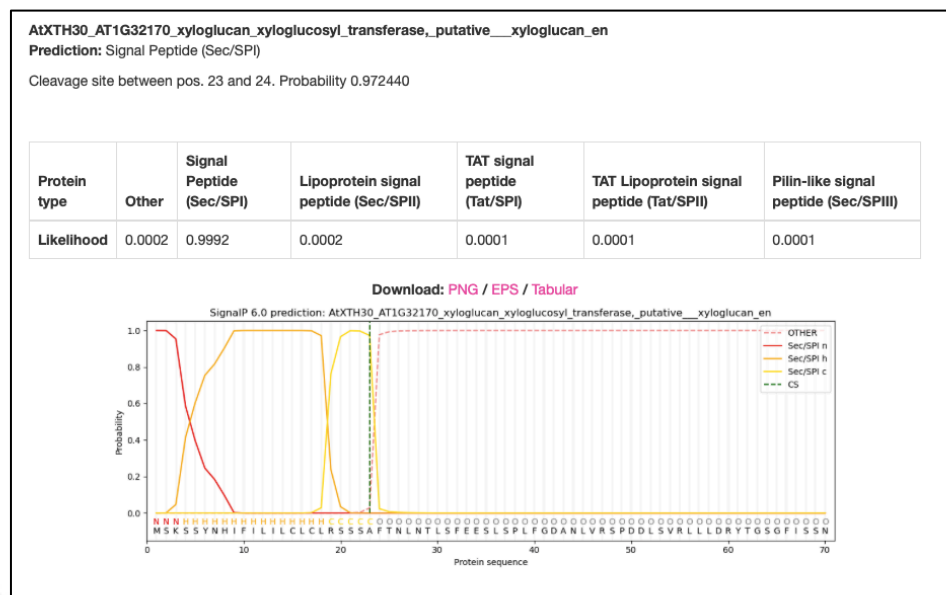


Figure 4.13. Predicted Signal Peptide and Cleavage Site of *AtXTH30*

Further investigation is based on N-linked glycosylation sites in selected XTH proteins. NetNGlyc 1.0 Server online tool, by searching for NXS/T in the sequences, uncovered that for *Equisetum*; 23 out of 25 sequences include an N-linked glycosylation motif. In *EhXTH_2007998* (*JVSZ_2007998*) and *EdXTH_2038789* (*CAPN_2038789*) there are no predicted N-linked glycosylation sites and may not contain signal peptides. There are sequences examined by NetNGlyc that contain N-linked glycosylation motif but may not contain signal peptide; *EhXTH_2009822*, *EhXTH_2008004*, *EhXTH_2128529*, *EdXTH_2039437*.

For *Arabidopsis thaliana*, 30 out of 33 sequences include N-linked glycosylation motifs. In contrast with SignalP, for *Arabidopsis thaliana*; in *AtXTH2*, *AtXTH31* and *AtXTH32* there are no detected N-linked glycosylation sites in the XTH sequences result of NetNGlyc analysis. For *AtXTH29*, even though it includes a potential glycosylation motif, it may not contain a possible signal peptide depending on the results of NetNGlyc.

4.3.4. DREME (Discriminative Regular Expression Motif Elicitation)

DREME MEME Suite Software was used to discovering the conserved protein motifs in *Equisetum* and *Arabidopsis thaliana* XTH protein sequences, by comparing them with control sets. The top 10 hits of conserved regions were given in Table 4.1 including the e-values and positive values among all sequences.

Table 4.1. DREME Results for *Equisetum* XTHs

Motif	Logo	E-value	Positives
1	DFEF	2.0e-010	25/25
2	WFDP	7.1e-008	23/25
3	WAT	1.6e-006	23/25
4	QTN	1.4e-004	23/25
5	YDYC	4.9e-004	20/25
6	VTAYY	6.4e-004	18/25
7	APF	2.3e-003	19/25
8	REQR	2.2e-003	19/25
9	FHTY	2.5e-003	17/25
10	KLVP	2.5e-003	17/25

Table 4.2. DREME Results for *Arabidopsis thaliana* XTHs

Motif	Logo	E-value	Positives
1	EFLG	3.2e-012	31/33
2	WFDP	9.5e-011	31/33
3	WAT	1.1e-009	32/33
4	YBYC	2.6e-009	28/33
5	DEID	9.5e-006	23/33
6	PIR	4.0e-005	30/33
7	HTY	1.4e-004	21/33
8	AGTV	1.6e-003	19/33
9	GSGF	1.6e-003	23/33
10	APF	3.0e-00.3	24/33

4.3.5. Protein Structure Homology Modeling

Modeling of protein structure of AtXTH30 carried out by SWISS-MODEL and Swiss PDB Viewer. Determination of highly conserved protein motifs by DREME analysis enlightens the protein modeling by checking the location of conserved motifs within the sequences by using SWISS-MODEL. The cartoon structure and surface model of AtXTH30 were modeled (Figure 4.14.). The catalytic motif “DEIDFEFLG” is colored in green (Figure 4.16.) and generated by The Bio-Analytic Resource for Plant Biology (<http://bar.utoronto.ca>).

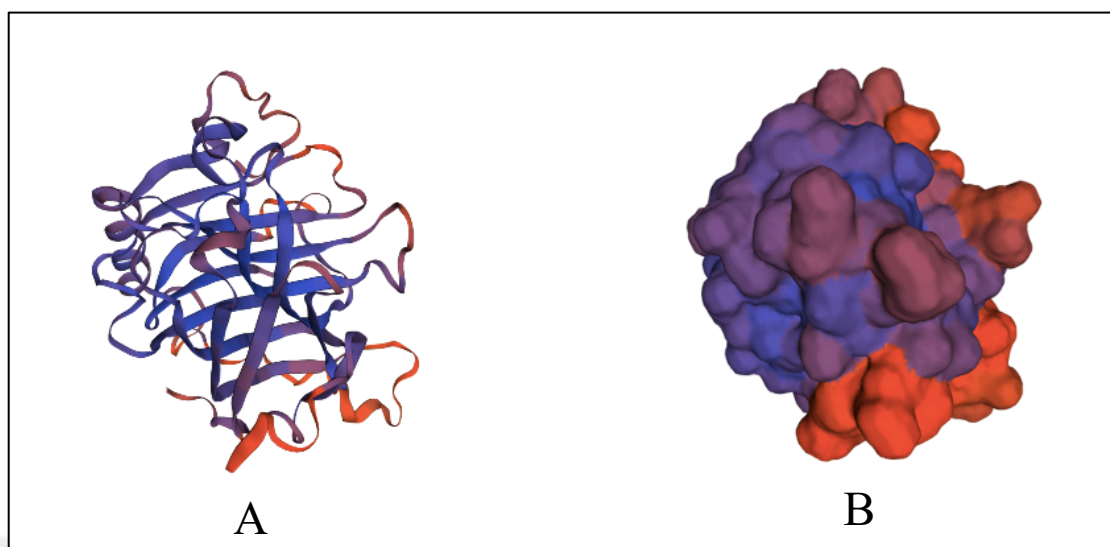


Figure 4.14. Predicted protein structure of *AtXTH30* by SWISS-MODEL.

A) Cartoon structure of *AtXTH30* created using 1UMZ.1 A as a template. B) Surface model of *AtXTH30* created using 1UMZ.1 A as a template.

The highly conserved regions are mapped on protein structure to determine the distance between the region and the substrate binding cleft and to check where they are. The aromatic residues within the XTH proteins, which are phenylalanine (F), tyrosine (Y), and tryptophan (W) amino acids, were examined by the Swiss PDB Viewer and colored in blue in surface model of *AtXTH30*, created by using 1.UMZ.1 A template (Figure 4.15.)

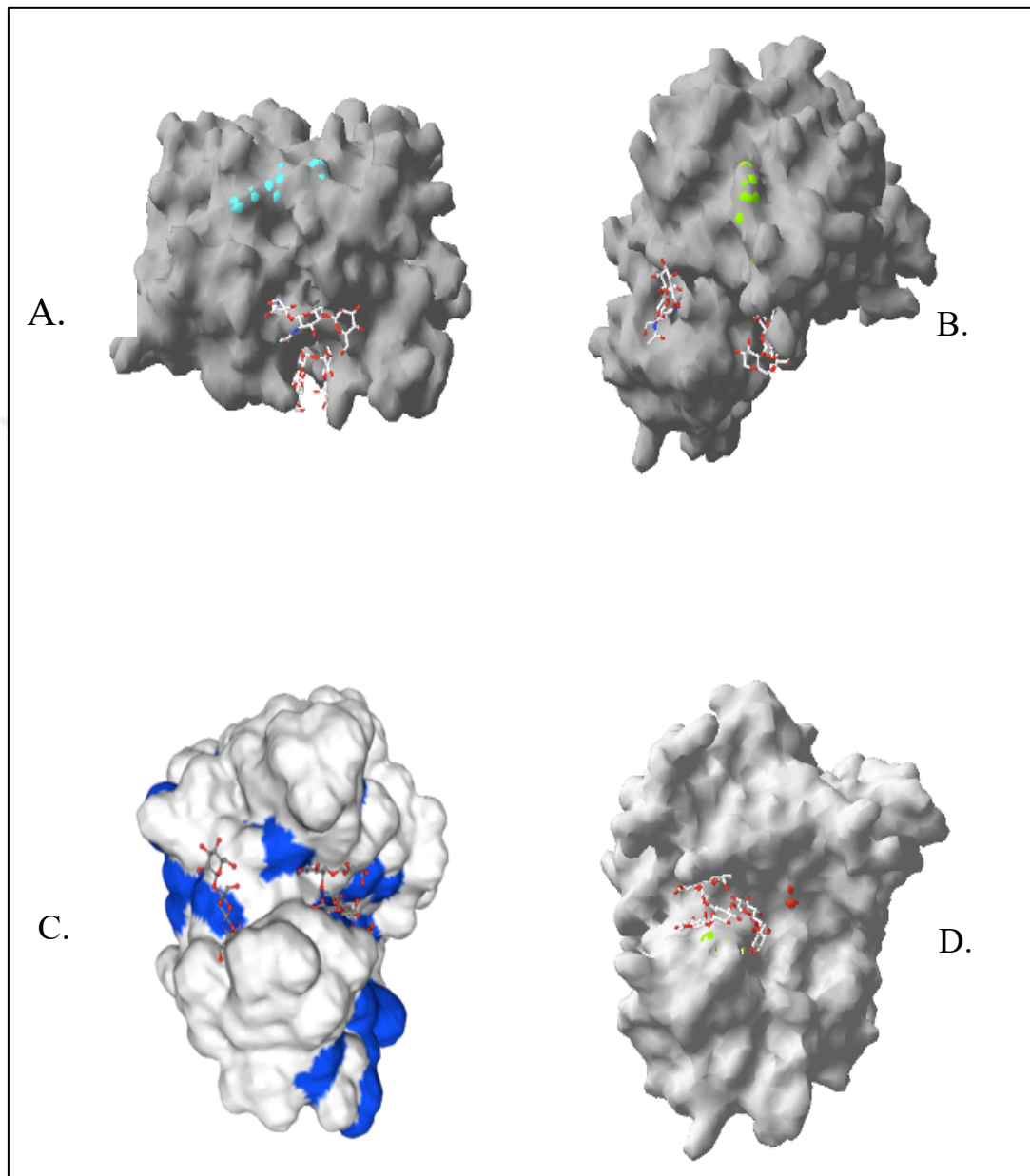


Figure 4.15. Predicted protein structure of *AtXTH30* with labeled conserved residues. A) Conserved region WFDP colored in light blue. B) Conserved region YBYC colored in green. C) Blue regions indicate aromatic residues F, Y, W. The left branches indicate the N-linked glycosylation, right branches indicates the XLLG substrate. D) Red region indicates the glutamic acids (E) of "DEIDFEFLG", light green region indicates WAT

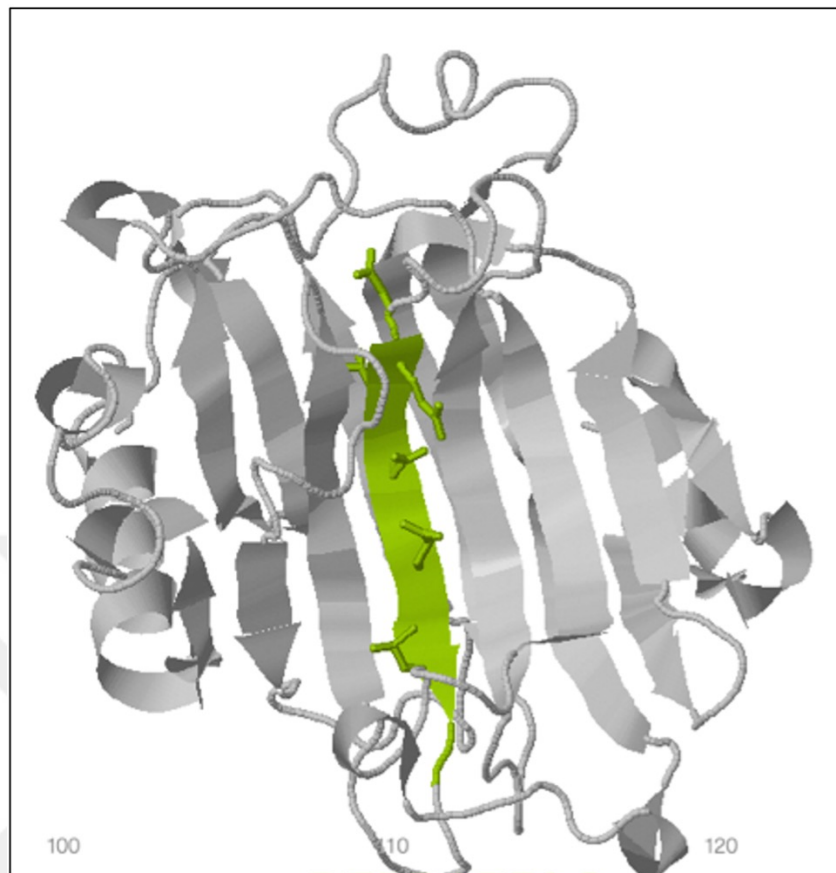


Figure 4.16. Predicted protein structure of *AtXTH30* with the green-colored catalytic site

5. DISCUSSION

The plant cell wall is an extremely crucial component that maintains plant cells together, and is a fundamental characteristic property of plant cells. Cell walls vary in structure and composition through the Plantae, but one feature they all have in common is that they all share the same component known as cellulose microfibrils encased in a matrix of structural proteins [65]. The structure of microfibrils inside narrow, extensible primary cell walls is dynamic, enabling cell growth in expanding tissues [66]. Cellulose biosynthesis is integrated with matrix polysaccharides in dicotyledonous plants' primary walls, including xyloglucans [67]. Numerous proteins and polysaccharides constitute the plant cell wall and, XyG is mainly composed of β -(1,4)-linked glucosyl residues that are joined together with xylosyl, fucosyl, and galactosyl residues [68]. Xyloglucan is an efficiently described hemicellulosic polysaccharide found in certain cotyledon plants' cell walls. [27]

For wall strength, the network formed by xyloglucan and cellulose microfibrils is critical. On the other hand, enzymes required for the catalyzing synthesis of xyloglucan, further those engaged in the alterations of xyloglucan and other wall polysaccharides, have yet to be identified. As a result, XTHs have been thoroughly investigated to elucidate the modifications to the cell wall structure. Alterations on xyloglucan cross-links were carried out by xyloglucan endotransglycosylase/hydrolases (XTHs) mostly important for the expansion and loosening of the plant cell wall. Accordingly, donor polysaccharides were cleaved by XTHs and resulted in relocating the reducing end of the original substrate to the oligosaccharide acceptor substrate or the nonreducing end of another substrate [69]. Furthermore, certain XTHs can accelerate the hydrolysis of donor polysaccharides.

This thesis attempted to heterologously express the *EtXTH1* gene from *Equisetum telmateia* and the *AtXTH30* gene from *Arabidopsis thaliana* in the methylotrophic yeast for active enzyme production by inducing the *AOXI* promoter and then characterization. Phylogenetic relationships of *Equisetum* and *Arabidopsis* species were examined separately to reveal the amino acid similarities and relationships between the amino acid sequences and the arrangement and composition of the plants. When comparing those amino acid similarities, XTH proteins are divided into six groups; Group I, Group II, Group III-A, Group III-B, EG16, and Group IV.

The products of the *XTH* gene belong to the genus of glycoside hydrolase family 16 (GH16), in addition, it includes a variety of microbial GH enzymes such as endoglucanases and chitin transglycosylases from the cell wall of *S. cerevisiae* [53]. Prior phylogenetic investigations of XTH gene products from the entire family have revealed that the bacterial licheninases, that preferentially hydrolyze linear mixed linkage β (1 $\rightarrow$ 3)/ β (1 $\rightarrow$ 4)-glucans; are the direct ancestors of the gene products of the xyloglucan endotransglycosylase/hydrolases [53]. XTH gene products are discernible from licheninases and such GH16 enzyme subfamilies by cleft-shaped catalytically active loop variations and a significant C-terminal extension, besides their selectivity for highly branched xyloglucan polysaccharides. Due to the remained ambiguous situation of molecular evolution of XTH gene products in GH16, biochemical and molecular investigations have been carried out to specifying the functions of endoglucanase which ended up with revealed the previously unidentified subfamily of GH16 participants transitional between bacterial licheninases and plant XTH gene products. Therefore, studies based on these hypotheses shed fresh light on the evolution of GH16 and bring together the seemingly diverse members of this essential protein family.

At the end of the molecular studies, in SDS-PAGE, protein bands were not sufficient to examine by their electrical charge. Therefore, western blotting was performed to identify proteins and recognize if there is any 6xHis-tag. The first large-scale production trial was successful but then identified proteins claimed that the expression target *AtXTH30* was not included in the yeast genome. Otherwise, it could be the reason that the alcohol oxidase I (AOX1) promoter, which drives protein production in yeast, was not induced with methanol in pPICZ α -C during the induction step [70]. On the contrary, the problem may have been caused at the beginning; bacterial transformation. If ligation of expression target into the expression vector was unsuccessful which means double-strand breaks of XTH genes were not enzymatically catalyzed by DNA ligases correctly, despite the optimal reaction conditions and replicate tests, it could cause breaks of the expression vector to ligate itself and transferred into *E.coli*.

Investigation of XTH enzymes is crucial for enlightening their substrate specificities and configurations of active sites, DEIDFEFLG or DEIDIEFLG, which is crucial for the enzymatic activity. Enzyme activity assays at various time intervals were supposed to show the donor polysaccharides and acceptor oligosaccharides but unfortunately, because

HPLC results did not give any donor-acceptor interaction, we could not proceed to measure enzyme activity. Determination of substrate specificities leads to modification of the substrate binding properties of higher plant XTH genes to improve the actions of XTHs in plant cell walls. Unfortunately, because we did not prove the presence of heterologously expressed XTH protein, TXG-XGO enzyme [71] activity tests were not carried out to see whether heterologously expressed or possess any XTH activity.

Due to examinations in multiple sequence alignments, we examined that catalytic site “DEIDFEFLG”, mostly the acidic amino acid glutamic acid (E), conserved among most XTHs with specific variations in plant genome. We also concluded by multiple sequence alignment that although there are amino acid replacements on XTH enzymes from *Equisetum* samples, substituted amino acids possess similar/identical characteristics with the amino acids that are present in the original catalytic motif [61] in most circumstances. For instance, *AtXTH30* has a catalytic motif “DELDIEFLG”, at the third amino acid instead of isoleucine (I) there is leucine (L) which both are nonpolar and aliphatic, and at the fifth amino acid instead of phenylalanine (F), there is isoleucine (I) which are just like the previous example those are nonpolar and hydrophobic amino acids.

In evolutionary analyses, the catalytic motif “DEIDFEFLG” that correlates with the activity of the XTH enzymes, has been found in both two target genes, *EtXTH1* and *AtXTH30*. Despite phylogenetic grouping and differences in potential catalytic activity, all *Equisetum* and *A. thaliana* XTH genes analyzed at the transcriptomic level share the conserved active site, DEIDFEFLG, located within the β - sheets. A higher plant that encodes XTH enzymes, *Arabidopsis thaliana*, has 33 genes for the XTH enzyme. Different land plants, including vascular and non-vascular to monocots and dicots, have approximately 20 to 60 genes [46]. The prevalence of these genes makes it obvious that XTH enzymes exhibit tissue-specificity [47] and particular characteristics. Depending on multiple environmental and developmental conditions, they show various expression levels [62]. Based on the generated phylogenetic tree, there is no *A. thaliana* XTHs were seen in the EG16 group, known as Endo-glucanase 16 Group thought to be evolved from bacterial licheninase enzymes [53].

All studied *Equisetum* members; *E. hymale*, *E. diffusum* and *E. telmateia* had EG16 Group enzymes. In Group IV, we had *AtXTH1*, *AtXTH2*, *AtXTH3*, *AtXTH11* and *OsXTH30*. *Equisetum* enzymes did not take part in Group IV, a recent study by Seven *et. al.* in 2021,

revealed that this group of enzymes is highly expressed in floral tissues [72], and due to *Equisetum* being a non-flowering plant, it is unlikely to be seen in Group IV. Another finding is, based on the phylogenetic tree, EG16 enzymes are closer to Group III-A and Group III-B enzymes. Each of the studied *Equisetum*; *E. hymale*, *E. diffusum* and *E. telmateia* have a member in the EG16 Group, which is significant because *Equisetum* is ancestral plants [73], also we had OsXTH31 in EG16. Due to generated phylogenetic tree, there are no *Arabidopsis* XTH genes belonging to the EG16 Group (Figure 4.9.), therefore it is suggested that *Arabidopsis thaliana* do not employ EG16 enzymes which have hydrolase activity, but AtXTH genes in Group III-A, could take the hydrolases role of the EG16's. All XTH sequences are assumed to have transferase activity, except for the EG16 group, which have hydrolase activity, and Group IIIA, which mostly have hydrolase activity with transferase activity.

All members of the AtXTHs shared an evolutionarily conserved active site (E*D*E) that contained the residues required for catalytic activity. The first glutamate residue (E), which catalyzes the enzymatic reaction, acts as a nucleophile, while the second glutamate residue, which acts as a base, activates the substrate. Each representative of the AtXTH family, except for AtXTH11 from Group IV, had a preserved N-glycosylation site that is close to the catalytic domains, as revealed. It is thought that the N-glycosylation site region contributes to the stability [74] of proteins. Mannosylated protein glycoforms from *AtXTH30* were created by heterologous expression in *Pichia pastoris*, however, they did not influence the enzyme's ability to function. [47]

Conservation of both glutamic acid (E) residues common base employed with b-glycosyl linkage cleavage [17]. The latest works verified the presence and functional significance of these residues in XTHs, as the breakdown of those linkages is the preliminary step of endotransglycosylase and endohydrolase functions [17]. Conserved amino acid residues inside the protein domains are actively engaged in vital activities, including the coordinating of enzyme catalytic sites. Therefore, at first, catalytic motif "DEIDFEFLG" and variations were predicted by DREME Software.

The identified motifs that were placed, including four aromatic residues, assumed to have a significant role in substrate binding. We examined; that most sequences contain a WAT region which is close to the highly conserved enzyme active site, and this tryptophan (W) is likely to be involved in substrate binding. Examination of the 3D structure of *A. thaliana*

XTHs, the positions of most highly conserved motifs, are located close to the substrate binding cleft. However, other highly conserved motifs such as WFDP and YBYC are located away from the substrate binding cleft.

The amount and position of predicted glycosylation sites in the transcribed *Equisetum* XTH sequences differed considerably. *EdXTH_2038789* and *EhXTH_2007998*, exhibited no identifiable glycosylation patterns, while other proteins exhibited up to three predicted sites. In addition, *EdXTH_2038789* and *EhXTH_2007998*, which lack any detectable N-glycosylation sites, are both putative hydrolases of the EG16 group. The evolutionarily conserved N-glycosylation site located directly after the enzymatic site is found in 23 of the encoded *Equisetum* XTH proteins and found in 30 of the Arabidopsis XTH proteins containing the identical region. Since glycosylation is extremely significant for the stability [74] of XTH proteins in heterologous expression studies, the absence of N-glycosylation might imply that all these enzymes possess short half-lives. The perception of that one, particularly for Group III-B enzymes such as *AtXTH30*, which is expected to have primarily hydrolase activity, could be critical to avoid excessive degradation. Four of the *Equisetum* XTH proteins *EhXTH_2009822*, *EhXTH_2008004*, *EhXTH_2128529*, and *EdXTH_2039437* might not contain signal peptides, interestingly all four of the proteins belong to Group II, known as having narrow substrate specificity, mostly prefer xyloglucan as a substrate in binding, were studied by Hrmova *et.al.* in 2002.

6. CONCLUSION

Xyloglucan endotransglycosylase/hydrolases (XTHs) enzymes are crucial for cell wall expansion and loosening related to the modifications of cross-linking of xyloglucan with cellulose and the network of xyloglucan itself. The plant cell wall that consists of active XTH enzymes is the most crucial segment that defense the plant against interactions and other environmental factors. Defining the structures, activities, and carbohydrate, and protein interactions in the plant cell wall is critical for improved yields, preserving the plant from stresses, and altering plant characteristics. The XTH genes found in entirely land plants belong to a diverse gene family with multiple branches on a dendrogram, and their functions in plants are ambiguous. All plant genomes examined thus far have a large number of genes for an enzyme that is meant to act on a single substrate.

Kinetic studies and substrate specificities based on plant cell-wall enzymes especially XTHs may be useful in plant selection for breeders. It will also reveal valuable insights on how to manipulate the current higher plant XTH genes to adjust substrate binding characteristics. Enlighten the evolutionary route of enzyme networks based on detecting the donor-acceptor substrate couples' specificities and enzyme activities. Multiple sequence alignments demonstrated that 33 XTH genes from *Arabidopsis thaliana* conserve the catalytic site "DEIDFEFLG" with mostly minor alterations while capable of having XTH activity. The cell wall modifications will be revealed once the roles of XTH enzymes are completely identified thus unsparing transgenic plants with gained properties help sustain in unfavorable environments. Furthermore, numerous innovations can be carried out in various fields – for example, papermaking, food science as texture and quality, brewing, etc.

Findings of this thesis may lead to the molecular phylogeny of XTH enzymes from different plant species with cell walls mostly containing pectic polysaccharides; for the growth and development of the plant, robustness, cell adhesion, and plant defense response coordination throughout this growth and development. This might contribute to further consumer-friendly cis-genic strategies. Hypothetically, one enzyme operates on a single form of xyloglucan substrate within cell wall, the presence of several xyloglucan structures requires the use of numerous XTH genes. The characterization of an XTH enzyme of *Equisetum telmateia* and *Arabidopsis thaliana* plants, including determining the molecular

phylogeny of XTH genes, was studied in this thesis to get a deeper insight into how those genes work. Hence plant growth and development pathways are complicated, the results of this thesis are insufficient to define if the targeted XTH enzymes are suitable for production or whether they are part of a more complex system. For transcriptome analysis of samples from *Equisetum* genus and obtained phylogenetic tree has shown that 33 *AtXTH* sequences, 31 *OsXTH*, and 67 *Equisetum* XTH protein sequences; 23 from *E. hymale*, 22 from *E. diffusum*, and 22 from *E. telmateia* identified here, all were detectable by having the highly conserved DEIDFEFLG motif that acts as the enzyme active site. Bioinformatic analyses of the *Equisetum* family and genome databases for XTH coding regions revealed the presence of 31 near-full lengths putative *XTH* sequences for *Equisetum*. By using DREME software, conserved regions were revealed, and by using SWISS-MODEL, *AtXTH30* were modeled by using 1umz.1 as a template. The generated phylogenetic tree revealed that; there are no *Equisetum* XTH genes belonging to Group IV and there are no *Arabidopsis* XTH genes belonging to EG16 Group.

For future perspectives, the significance of XTH enzymes in other plants, along with the expression of the above-mentioned genes in diverse tissues of varied plants, should be studied in future studies to better comprehend their mechanisms and networks. To learn more about the enzyme structure and enzyme-substrate interactions that are related to a determination of the amino acids that affect the binding or catalysis, looking into molecular modeling, if there are at least two molecules molecular docking, and protein crystallization methods could be useful. Molecular modeling refers to the visualization and manipulation of 3-D structures, including the estimation of physicochemical characteristics that may assist in defining the structural activity relationship abbreviated as SAR, in this case, *AtXTH30*. Therefore, molecular modeling of *AtXTH30* provides structure representation with C-terminal extension, beta-sheets, and highlighted catalytic amino acids. In this thesis, the partial sequences of *Equisetum* genus, *Arabidopsis thaliana*, and *Oryza sativa* were used for multiple sequence alignments. For further investigations, NGS-based whole genome sequencing (WGS) could be used to obtain the most comprehensive data about target organisms. Also, hydrophobic cluster analysis may be performed further with data obtained from multiple sequence alignments on ClustalW to investigate the secondary structure distribution.

REFERENCES

1. Koornneef M, Meinke D. The development of Arabidopsis as a model plant. *Plant Journal*. 2010;61:909–921.
2. Boyes DC, Zayed A, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, *et al.* Growth stage-based phenotypic analysis of Arabidopsis: A model for high throughput functional genomics in plants. *Plant Cell*. 2001;13:1499–1510.
3. Wang NN, Li Y, Miao M, Zhu CH, Yan KJ, Jiang ZQ. High level expression of a xyloglucanase from *Rhizomucor miehei* in *Pichia pastoris* for production of xyloglucan oligosaccharides and its application in yoghurt. *International Journal of Biological Macromolecules*. 2021;190:845–852.
4. Yokoyama R, Nishitani K. A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of Arabidopsis. *Plant and Cell Physiology*. 2001; 42:1025–1033.
5. González-Rábade N, Del Carmen Oliver-Salvador M, Salgado-Manjarrez E, Badillo, Corona JA. In vitro production of plant peroxidases - A review. *Applied Biochemistry and Biotechnology*. 2012;166:1644–1660.
6. Becnel J, Natarajan M, Kipp A, Braam J. Developmental expression patterns of Arabidopsis XTH genes reported by transgenes and Genevestigator. *Plant Molecular Biology*. 2006;61:451–467.
7. Krämer U. Planting molecular functions in an ecological context with Arabidopsis thaliana. *eLife*. 2015;4:1–13.
8. Zhang B, Gao Y, Zhang L, Zhou Y. The plant cell wall: Biosynthesis, construction, and functions. *Journal of Integrative Plant Biology*. 2021;63:251–272.
9. Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA. A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant Journal*. 2016;88:1058–1070.

10. Radojević ID, Stanković MS, Stefanović OD, Topuzović MD, Čomić LR, Ostojić AM. Great horsetail (*Equisetum telmateia* Ehrh.): Active substances content and biological effects. *EXCLI Journal*. 2012;11:59–67.
11. Law C, Exley C. New insight into silica deposition in horsetail (*Equisetum arvense*). *BMC Plant Biology*. 2011;11.
12. Yeganegi M, Yazdi F, Mortazavi S, Asili J. *Equisetum telmateia* extracts: Chemical compositions, antioxidant activity and antimicrobial effect on the growth of some pathogenic strain causing poisoning and infection. *Microbial Pathogenesis*. 2018.
13. Gu Y, Rasmussen CG. Cell biology of primary cell wall synthesis in plants. *Plant Cell*. 2022;34:103–128.
14. Schenk HJ, Jansen S, Hölttä T. Positive pressure in xylem and its role in hydraulic function. *New Phytologist*. 2021;230:27–45.
15. Knox JP. Developmentally glycoproteins regulated of the plant proteoglycans cell surface and. *The FASEB Journal*. 1995;9:1004–1012.
16. Gür B. Multi-omics Approach To Investigate The Roles Of Xyloglucan Endotransglycosylase/Hydrolases (XTHs) In Agriculturally Important Plants, Doctoral Thesis, Yeditepe University, 2018.
17. Van Sandt VST, Guisez Y, Verbelen JP, Vissenberg K. Analysis of a xyloglucan endotransglycosylase/hydrolase (XTH) from the lycopodiophyte *Selaginella kraussiana* suggests that XTH sequence characteristics and function are highly conserved during the evolution of vascular plants. *Journal of Experimental Botany*. 2006;57:2909–2922.
18. Shinohara N, Sunagawa N, Tamura S, Yokoyama R, Ueda M, Igarashi K, *et al*. The plant cell-wall enzyme AtXTH3 catalyses covalent cross-linking between cellulose and cello-oligosaccharide. *Scientific Reports*. 2017;7:1–10.
19. Elgorriaga A, Escapa I, Rothwell G, Tomescu AMF, Ruben Cuneo N. Origin of *Equisetum*: Evolution of horsetails (*Equisetales*) within the major euphyllophyte clade *Sphenopsida*. *American Journal of Botany*. 2018;105: 1286–1303.

20. Lampugnani ER, Khan, GA, Somssich, M, Persson S, *et al.* Building a plant cell wall at a glance. *Journal of Cell Science*. 2018;131(2):jcs207373.
21. Ochoa-Villarreal M, Aispuro-Hernández E, Vargas-Arispuro I, Ngel M. Plant Cell Wall Polymers: Function, Structure and Biological Activity of Their Derivatives. *Polymerization*. 2012.
22. Brett CT. Cellulose microfibrils in plants: Biosynthesis, deposition, and integration into the cell wall. *International Review of Cytology*. 2000;199:161–199.
23. Lerouxel O, Cavalier DM, Liepman AH, Keegstra K. Biosynthesis of plant cell wall polysaccharides - a complex process. *Current Opinion in Plant Biology*. 2006;9:621–630.
24. Cosgrove DJ. Growth of the plant cell wall. *Nature Reviews Molecular Cell Biology*. 2005;6:850–861.
25. Chami Khazraji A, Robert S. Self-assembly and intermolecular forces when cellulose and water interact using molecular modeling. *Journal of Nanomaterials*. 2013;2013.
26. Levy S, MacLachlan G, Staehelin LA. Xyloglucan sidechains modulate binding to cellulose during in vitro binding assays as predicted by conformational dynamics simulations. *Plant Journal*. 1997;11:373–386.
27. Zabortina OA. Xyloglucan and its biosynthesis. *Frontiers in Plant Science*. 2012;3:2008–2013.
28. Schultink A, Liu L, Zhu L, Pauly M. Structural diversity and function of xyloglucan sidechain substituents. *Plants*. 2014;3:526–542.
29. Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink, A, *et al.* Hemicellulose biosynthesis. *Planta*. 2013;238:627–642.
30. Ravachol J, De Philip P, Borne R, Mansuelle P, Maté MJ, Perret S, Fierobe HP *et al.* Mechanisms involved in xyloglucan catabolism by the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. *Scientific Reports*. 2016;6:1–17.

31. Cavalier DM, Lerouxel O, Neumetzler L, Yamauchi K, Reinecke A, Freshour G, *et al.* Disrupting two *Arabidopsis thaliana* xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell*. 2008;20:1519–1537.
32. Rose JKC, Braam J, Fry SC, Nishitani K. The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. *Plant and Cell Physiology*. 2002;43:1421–1435.
33. Chen Z, Li S, Fu Y, Li C, Chen D, Chen, H. Arabinoxylan structural characteristics, interaction with gut microbiota and potential health functions. *Journal of Functional Foods*. 2019;54:536–551.
34. York WS, O'Neill MA. Biochemical control of xylan biosynthesis - which end is up? *Current Opinion in Plant Biology*. 2008;11:258–265.
35. Popper ZA, Fry SC. Xyloglucan-pectin linkages are formed intra-protoplasmically, contribute to wall-assembly, and remain stable in the cell wall. *Planta*. 2008;227:781–794.
36. Xue X, Fry SC. Evolution of mixed-linkage (1,3, 1,4)- β -D-glucan (MLG) and xyloglucan in *Equisetum* (horsetails) and other monilophytes. *Annals of Botany*. 2012;109:873–886.
37. Ratke C, Terebieniec, Barbara K, Winstrand S, Derba-Maceluch M, Grahn T, Schiffthaler B, *et al.* Downregulating aspen xylan biosynthetic GT43 genes in developing wood stimulates growth via reprogramming of the transcriptome. *New Phytologist*. 2018;219:230–245.
38. Jensen JK, Busse-Wicher M, Poulsen CP, Fangel JU, Smith P J, Yang JY *et al.* Identification of an algal xylan synthase indicates that there is functional orthology between algal and plant cell wall biosynthesis. *New Phytologist*. 2018;218:1049–1060.
39. Chong SL, Virkki L, Maaheimo H, Juvonen M, Derba-Maceluch M, Koutaniemi S, Roach M, *et al.* O-Acetylation of glucuronoxylan in *Arabidopsis thaliana* wild type and its change in xylan biosynthesis mutants. *Glycobiology*. 2014;24:494–506.

40. Mortimer JC, Faria-Blanc N, Yu X, Tryfona T, Sorieul M, Ng YZ, *et al.* An unusual xylan in Arabidopsis primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *Plant Journal*. 2015;83:413–426.
41. Handford MG, Baldwin TC, Goubet F, Prime TA, Miles J, Yu X, *et al.* Localisation and characterisation of cell wall mannan polysaccharides in Arabidopsis thaliana. *Planta*. 2003;218:27–36.
42. Davis J, Brandizzi F, Liepman AH, Keegstra K. Arabidopsis mannan synthase CSLA9 and glucan synthase CSLC4 have opposite orientations in the Golgi membrane. *Plant Journal*. 2010;64:1028–1037.
43. Driouich A, Follet-Gueye ML, Bernard S, Kousar S, Chevalier L, Vicré-Gibouin M, *et al.* Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Frontiers in Plant Science*. 2012;3:1–15.
44. Ström A, Ribelles P, Lundin L, Norton I, Morris ER, Williams MAK. Influence of pectin fine structure on the mechanical properties of calcium-pectin and acid-pectin gels. *Biomacromolecules*. 2007;8:2668–2674.
45. Glinsky VV., Raz A. Modified citrus pectin anti-metastatic properties: one bullet, multiple targets. *Carbohydrate Research*. 2009;344:1788–1791.
46. Eklöf JM, Brumer H. The XTH gene family: An update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiology*. 2010;153:456–466.
47. Fu MM, Liu C, Wu F. Genome-Wide Identification, Characterization and Expression Analysis of Xyloglucan Endotransglucosylase/Hydrolase Genes Family in Barley (*Hordeum vulgare*). *Molecules*. 2019;24(10):1935.
48. Rose JK, Braam J, Fry SC, Nishitani K. The XTH family of enzymes involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiol*. 2002;43(12):1421–1435.
49. Caffall KH, Mohnen D. The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydrate Research*. 2009;344:1879–1900.

50. Juturu V, Wu JC. Heterologous Protein Expression in *Pichia pastoris*: Latest Research Progress and Applications. *ChemBioChem*. 2018;19:7–21.
51. Kaewthai N, Harvey AJ, Hrmova M, Brumer H, Ezcurra I, Teeri TT. *et al.* Heterologous expression of diverse barley XTH genes in the yeast *Pichia pastoris*. *Plant Biotechnology*. 2010;27:251–258.
52. Karbalaei M, Rezaee SA, Farsiani H. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *Journal of Cellular Physiology*. 2020;235:5867–5881.
53. Eklöf JM, Shojania S, Okon M, McIntosh LP, Brumer, H. Structure-function analysis of a broad specificity *Populus trichocarpa* endo- β -glucanase reveals an evolutionary link between bacterial licheninases and plant XTH gene products. *Journal of Biological Chemistry*. 2013;288:15786–15799.
54. Baumann MJ, Eklöf JM, Michel G, Kallas ÅM, Teeri TT, Czjzek M, *et al.* Structural evidence for the evolution of xyloglucanase activity from xyloglucan Endo-transglycosylases: Biological implications for cell wall metabolism. *Plant Cell*. 2007;19:1947–1963.
55. Nowakowski AB, Wobig WJ, Petering DH. Native SDS-PAGE: High resolution electrophoretic separation of proteins with retention of native properties including bound metal ions. *Metallomics*. 2014;6:1068–1078.
56. Sanchez L. TCA protein precipitation protocol. [cited 2022 20 June] Available from: http://www.its.caltech.edu/~bjorker/TCA_ppt_protocol.pdf
57. Seven M. Characterization of Xyloglucan Endotransglycosylase (XET) Enzymes From Different Plant Species. Doctoral Thesis, Yeditepe University 2017.
58. Li P, Anumanthan A, Gao XG, Ilangoan K, Suzara VV, Düzgüneş, N, *et al.* Expression of recombinant proteins in *Pichia pastoris*. *Applied Biochemistry and Biotechnology*. 2007;142:105–124.

59. Lin-Cereghino J, Hashimoto MD, Moy A, Castelo J, Orazem CC, Kuo P, *et al.* Direct selection of *Pichia pastoris* expression strains using new G418 resistance vectors. *Yeast*. 2008;25:293–299.
60. Hu X, Mao R, Zhang Y, Teng D, Wang X, Xi D, *et al.* Biotechnical paving of recombinant enterocin A as the candidate of anti-*Listeria* agent. *BMC Microbiology*. 2014;14:1–11.
61. Akdemir H, Seven M, Derman C, Harvey AJ. In silico analysis of XTH gene family from barley (*Hordeum vulgare* L.) and their comparative expression analysis during germination. *Turkish Journal of Botany*. 2022;46:92–108.
62. Plackett ARG, Powers SJ, Phillips AL, Wilson ZA, Hedden P, Thomas SG. The early inflorescence of *Arabidopsis thaliana* demonstrates positional effects in floral organ growth and meristem patterning. *Plant Reproduction*. 2018;31:171–191.
63. Smyth DR, Bowman JL, Meyerowitz EM. Early Flower Development in *Arabidopsis*. *The Plant Cell*. 1990;2:755–767.
64. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, *et al.* A gene expression map of *Arabidopsis thaliana* development. *Nature Genetics*. 2005;37:501–506.
65. Kotler P. Plant Cell Wall Patterning and Cell Shape. New Jersey: John Wiley & Sons Inc. 2021;2–5.
66. Polko JK, Kieber JJ. The regulation of cellulose biosynthesis in plants. *Plant Cell*. 2019;31:282–296.
67. Carpita NC, Gibeaut DM. Structural models of primary cell walls in flowering plants: Consistency of molecular structure with the physical properties of the walls during growth. *Plant Journal*. 1993;3:1–30.
68. Peña MJ, Darvill AG, Eberhard St, York WS, O'Neill MA. Moss and liverwort xyloglucans contain galacturonic acid and are structurally distinct from the xyloglucans synthesized by hornworts and vascular plants. *Glycobiology*. 2008;18:891–904.

69. Holland C, Simmons TJ, Meulewaeter F, Hudson A, Fry SC. Three highly acidic Equisetum XTHs differ from hetero-trans- β -glucanase in donor substrate specificity and are predominantly xyloglucan homo-transglucosylases. *Journal of Plant Physiology*. 2020;251:153210.
70. Chang CH, Hsiung HA, Hong KL, Huang CT. Enhancing the efficiency of the *Pichia pastoris* AOX1 promoter via the synthetic positive feedback circuit of transcription factor Mxr1. *BMC Biotechnology*. 2018;18:1–10.
71. Dünkelfmann P, Kolter-Jung D, Nitsche A, Demir AS, Siegert P, Lingen B, *et al.* Development of a donor-acceptor concept for enzymatic cross-coupling reactions of aldehydes: The first asymmetric cross-benzoin condensation. *Journal of the American Chemical Society*. 2002;124:12084–12085.
72. Seven M, Derman C, Harvey AJ. Enzymatic characterization of ancestral/group-IV clade xyloglucan endotransglycosylase/hydrolase enzymes reveals broad substrate specificities. *Plant Journal*. 2021;106:1660–1673.
73. Vanneste K, Sterck L, Myburg AA, Van de Peer Y, Mizrachi E. Horsetails are ancient polyploids: Evidence from *Equisetum giganteum*. *Plant Cell*. 2015;27:1567–1578.
74. Sarkar A, Wintrode PL. Effects of glycosylation on the stability and flexibility of a metastable protein: The human serpin α 1-antitrypsin. *International Journal of Mass Spectrometry*. 2011;302:69–75.
75. Clark JW, Puttick MN, Donoghue PCJ. Origin of horsetails and the role of whole-genome duplication in plant macroevolution. *Proceedings of the Royal Society B: Biological Sciences*. 2019;286.