

**Analyzing structure and function
of an industrially important enzyme from
Nectria haematococca for biofuel
production**

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

der Universität Hamburg

Fachbereich Chemie

vorgelegt von

Hina Andaleeb

aus Multan, Pakistan

Hamburg, 2021

Die vorliegende Arbeit wurde im Zeitraum von Mai 2017 bis November 2020 in der Arbeitsgruppe von Prof. Ch. Betzel im Laboratorium für Strukturbiologie von Infektion und Entzündung am DESY und am Institut für Biochemie und Molekularbiologie, des Fachbereichs Chemie der Universität Hamburg, durchgeführt.

1. Gutachter: Prof. Dr. Christian Betzel

2. Gutachter: Prof. Dr. Andrew Torda

Tag der Disputation: 23.04.2021

**I DEDICATE THIS WORK
TO MY LOVELY MOTHER, HUSBAND AND DAUGHTER**

Publications

Parts of this thesis have been published or are in the process of being published:

1. **Hina Andaleeb**, Najeeb Ullah, Sven Falke, Markus Perbandt, Hévila Brognaro & Christian Betzel. **2020**. High-resolution crystal structure and biochemical characterization of a GH11 endoxylanase from *Nectria haematococca*. *Scientific Reports*, *10* (1), 1-13.
2. Inhibition studies of GH11 endoxylanase from *Nectria haematococca* by lignin-derived phenolic compounds, in preparation.

The author has also contributed to the following publication:

3. Najeeb Ullah, **Hina Andaleeb**, Celestin Nzanzu Mudogo, Sven Falke, Christian Betzel and Carsten Wrenger. **2020**. Solution Structures and Dynamic Assembly of the 24-Meric Plasmodial Pdx1-Pdx2 Complex. *International Journal of Molecular Sciences*, *21* (17), 5971.

Table of contents

Publications.....	i
Table of contents	ii
List of Abbreviations	v
Abstract.....	ix
Zusammenfassung.....	xi
1. Introduction.....	1
1.1 Importance of lignocellulose biomass	1
1.2 Biofuels versus fossil fuels	2
1.3 <i>Nectria haematococca</i> : a source of enzymes for biofuel production	4
1.4 Optimum conditions of carbohydrate degrading enzymes	5
1.5 Industrial applications of xylanases.....	6
1.5.1 Food and feed technology.....	6
1.5.2 Xylooligosaccharides.....	8
1.5.3 Paper industry	8
1.5.4 Cleaning agents.....	8
1.5.5 Wine and juice industries.....	8
1.5.6 Biofuel	9
1.6 Introductory preface of X-ray crystallography.....	15
1.6.1 Conventional vs serial crystallography.....	17
1.6.2 Crystallographic studies of xylanases.....	19
1.7 Inhibition of xylanases	22
1.8 Scope of the study	24
2. Experimental part.....	26
2.1 General chemicals	26
2.1.1 Material and manufacturer.....	26
2.2 Mutations and truncation of <i>N. haematococca</i> xylanase (NhGH11)	29
2.3 Transformation of competent <i>E. coli</i> cells	30
2.4 Gene expression.....	31
2.5 Cell lysis and affinity chromatography	32
2.6 Protein quantification	34

2.7 Expression and purification of TEV protease	34
2.8 Sample concentration	36
2.9 Size exclusion chromatography (SEC).....	36
2.10 SDS polyacrylamide gel electrophoresis (SDS-PAGE).....	36
2.11 Dynamic light scattering.....	38
2.12 Circular dichroism (CD) spectroscopy	38
2.13 Crystallization.....	39
2.13.1 Diffraction data collection	40
2.13.2 Tape-drive experiments	41
2.13.3 Soaking experiments.....	41
2.14 Molecular docking.....	44
2.15 Biochemical characterization	44
2.15.1 Effect of temperature and pH on enzyme activity	44
2.15.2 Effect of metal ions and chemical reagents	45
2.15.3 Substrate kinetics	45
2.15.4 Identification of products.....	46
2.16 Inhibition assays	46
2.16.1 Residual activity	46
2.16.2 Thermofluor assays.....	48
2.16.3 Tryptophan fluorescence spectrophotometry	48
3. Results and discussion	50
3.1 Surface entropy reduction prediction (SERp)	50
3.2 Protein expression and purification	51
3.3 Crystallization experiments	53
3.3.1 Data collection at Synchrotron	54
3.3.2 Refinement of single crystal data	55
3.3.3 Crystallographic structure of NhGH11.....	58
3.3.4 Surface exposed aromatic amino acids.....	60
3.3.5 Catalytic cavity of NhGH11	61
3.3.6 Analysis of clefts and binding sites for sugars	63
3.4 Sequence homology.....	65
3.5 Elucidation of thermostability parameters.....	69
3.6 Tape drive experiments	75

3.7 Co-crystallization and soaking experiments.....	78
3.8 Biochemical and functional characterization of NhGH11	81
3.8.1 Optimum temperature	81
3.8.2 Optimum pH.....	82
3.8.3 Effect of metal ions and chemical reagents	84
3.8.4 Substrate specificity and kinetic parameters.....	87
3.8.5 Identification of products.....	90
3.9 Inhibition studies	92
3.9.1 Residual activity	92
3.9.2 Thermofluor assay	94
3.9.3 Tryptophan fluorescence spectrophotometry	99
3.10 Conclusion and outlook.....	105
4. Supplementary information	107
5. Bibliography	119
6. List of Figures.....	142
7. List of Tables	144
8. Risks and safety statements.....	145
9. Acknowledgment.....	152
10. Eidesstattliche Erklärung.....	154

List of Abbreviations

Å	Ångström
AA	Auxiliary activities
Approx.	Approximately
APS	Ammonium persulfate
C	Celsius
CAZy	Carbohydrate-active enzymes
CBH	Cellobiohydrolases
CD	Circular dichroism
Da	Dalton
3D	Three dimensional
ddH₂O	Double-distilled water
DESY	Deutsches elektronen synchrotron
DLS	Dynamic light scattering
DNS	Dinitrosalicylic acid
DTT	(2S,3S)-1,4-Bis(sulfanyl)butane-2,3-diol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	2,2',2'',2''' - (1,2-Ethanediyldinitrilo) tetraacetic acid
EMBL	European molecular biology laboratory
et al.	et alii/et aliae
e.g	Exempli gratia or for example
FPLC	Fast protein liquid chromatography
g	Gram
GH	Glycosyl hydrolase
h	Hours
i.e	id est or that is
IPTG	Isopropyl β- d-1-thiogalactopyranoside
k	Kilo
Km	Michaelis constant
L	Liter

LB	Luria bertani
LPMO	Lytic polysaccharide monooxygenase
M	Molar
m	Meter
MAD	Multi wavelength anomalous diffraction
MALDI	Matrix-assisted laser desorption ionization
MBP	Maltose binding protein
MIR	Multiple isomorphous replacement
min	Minutes
mm	Millimeter
MR	Molecular replacement
MS	Mass spectrometry
MW	Molecular weight
NTA	Nitrilotriacetic acid
NhGH11	<i>Nectria haematococca</i> xylanase from glycosyl hydrolase family 11
nm	nanometer
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank
PEG	Polyethylene glycol
PETRA	Positron elektron tandem ring anlage
pH	Power of hydrogen
R_H	Hydrodynamic radius
RMSD	Root mean square deviation
rpm	Rounds per minute
s	Seconds
SAD	Single-wave anomalous diffraction
SDS	Sodium dodecyl sulfate
SEC	Size exclusion chromatography
SERp	Surface entropy reduction prediction

SIR	Single isomorphous replacement
sp	Subspecies
TEV	Tobacco etch virus
TAXI	Triticum aestivum xylanase inhibitors
TEMED	Tetramethylethylenediamine
T_m	Melting temperature
TRIS	Tris- (hydroxymethyl)-aminomethane
v/v	Volume per volume
V_{max}	Maximum velocity
W	Watt
w/v	Weight per volume
α	Alpha
β	Beta
γ	Gamma
μ	Micro
X	Xylose
X₂	Xylobiose
X₃	Xylotriose
X₄	Xylotriose
X₅	Xylopentose
XIP	Xylanase inhibitor proteins
XOSs	Xylooligosaccharides

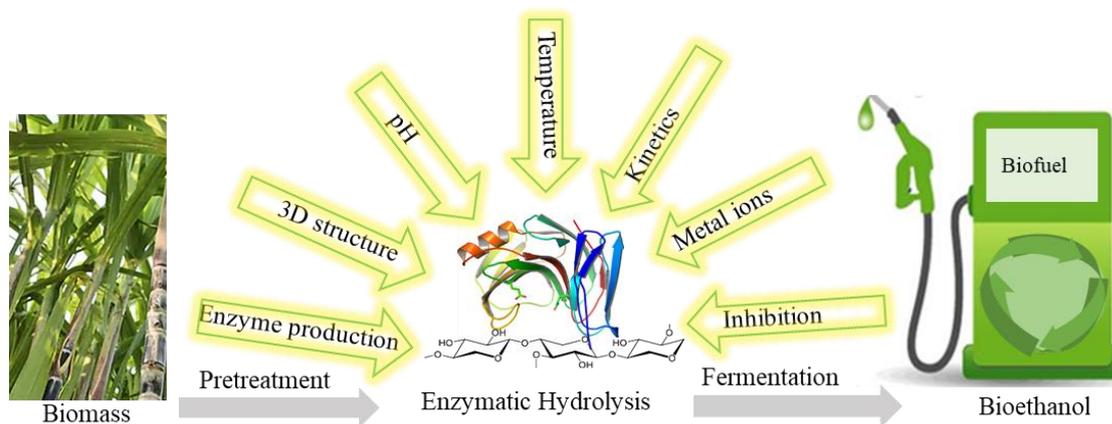
One letter code	Three letter code	Amino Acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine

List of Abbreviations

I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
T	Ser	Serine
S	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Abstract

The demand for biofuels is increasing constantly with industrialization and population growth. Biofuels can replace fossil fuels due to their effective role in reducing carbon dioxide emissions and global warming. Increasing demand for biofuels triggered substantial scientific research activities to optimize the efficiency of biofuel production. According to the annual oil consumption statistics of Pakistan, the remaining oil reserves can fulfill the requirements only for the next two years. Pakistan and other developing countries that have a rich source of agricultural biomass can apply the biofuel production technology to reduce or even replace the use of fossil fuels. Agricultural wastes, especially sugar cane bagasse, corn and wheat straws are potential lignocellulosic waste materials.



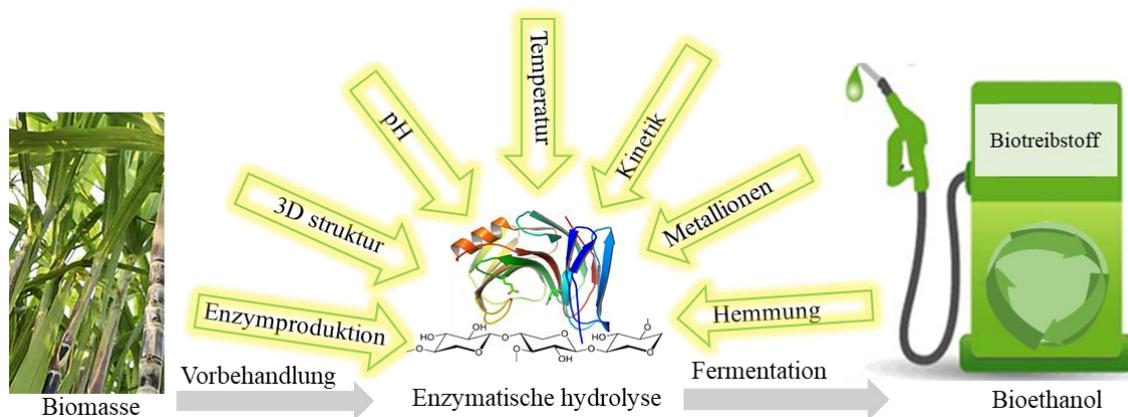
Referring to the carbohydrates active enzymes for biofuel production, the genome of *Nectria haematococca* revealed several putative biomass-degrading enzymes, which have been structurally uncharacterized so far, including the glycoside hydrolase from family 11. Therefore, xylanase from *N. haematococca* (NhGH11) was selected as a potential candidate for biophysical and biochemical characterization. The three-dimensional structure of xylanase from *N. haematococca* (NhGH11) is determined at high-resolution (1.0 Å). NhGH11 is single domain xylanase with β -jelly roll folds, the antiparallel and highly twisted sheets create a well-structured cylindrical cavity for the binding of xylan substrates. Due to the advantage of a high-resolution structure, NhGH11 provides details and insights about the complex hydrogen-bonding network of the active site region and allows a detailed comparison with homologous structures. Biochemical and functional properties show the catalytic potential of NhGH11 to be applied for industrial applications. NhGH11 can catalyze the hydrolysis of complex xylan into simple xylose subunits of

different lengths. The mesophilic character of NhGH11 is confirmed by the comparison of the number of hydrogen bonds, side-chain interactions and buried water molecules, with thermophilic GH11 enzymes. NhGH11 is not found to be sensitive towards metal ions and chemical reagents typically present in associated industrial production processes. However, lignin-derived phenolic compounds i.e. caffeic acid, cinnamic acid, coumaric acid, and ferulic acid can affect the catalysis by inhibiting the xylanase activity of NhGH11.

Therefore, this data highlights the potential of NhGH11 to be applied in industrial biomass degradation processes, in which the catalysis of xylan biopolymers can produce sustainable biofuel and bio-based commercial eco-friendly materials.

Zusammenfassung

Die Nachfrage nach Biokraftstoffen steigt mit der Industrialisierung und dem Bevölkerungswachstum stetig an. Biokraftstoffe können fossile Brennstoffe ersetzen, da sie die Kohlendioxidemissionen und die globale Erwärmung wirksam reduzieren. Die steigende Nachfrage nach Biokraftstoffen löste umfangreiche wissenschaftliche Forschungsaktivitäten aus, um die Effizienz der Biokraftstoffproduktion zu optimieren. Nach der jährlichen Ölverbrauchsstatistik Pakistans können die verbleibenden Ölreserven die Anforderungen nur für die nächsten zwei Jahre erfüllen. Pakistan und andere Entwicklungsländer mit einer reichen Quelle landwirtschaftlicher Biomasse können die Biokraftstoffproduktionstechnologie anwenden, um den Einsatz fossiler Brennstoffe zu reduzieren oder sogar zu ersetzen. Landwirtschaftliche Abfälle, insbesondere Zuckerrohrbagasse, Mais und Weizenstrohhalm, sind potenzielle Abfälle von Lignocellulose.



In Bezug auf die kohlenhydrataktiven Enzyme für die Biokraftstoffproduktion enthüllte das Genom von *Nectria haematococca* mehrere mutmaßliche biomasseabbauende Enzyme, die bisher strukturell nicht charakterisiert wurden, einschließlich der Glycosidhydrolase aus Familie 11. Daher wurde die Xylanase aus *N. haematococca* (NhGH11) als potenzieller Kandidat für die biophysikalische und biochemische Charakterisierung ausgewählt. Die dreidimensionale Struktur der Xylanase aus *N. haematococca* (NhGH11) wurde bei hoher Auflösung (1.0 Å) bestimmt. NhGH11 ist eine Einzeldomänen-Xylanase mit β -jelly roll-Faltung. Die antiparallelen und stark verdrillten Schichten bilden einen gut strukturierten zylindrischen Hohlraum für die Bindung von Xylansubstraten. Die hochaufgelöste Struktur von NhGH11 liefert Details und Einblicke in das komplexe Wasserstoffbindungsnetzwerk

der Region des aktiven Zentrums und ermöglicht einen detaillierten Vergleich mit homologen Strukturen. Biochemische und funktionelle Eigenschaften zeigen das katalytische Potenzial von NhGH11 für industrielle Anwendungen. NhGH11 kann die Hydrolyse von komplexem Xylan zu einfachen Xylose-Untereinheiten unterschiedlicher Länge katalysieren. Der mesophile Charakter von NhGH11 wird durch den Vergleich der Anzahl von Wasserstoffbrückenbindungen, Seitenkettenwechselwirkungen und verborgenen Wassermolekülen mit thermophilen GH11-Enzymen bestätigt. Es wurde festgestellt, dass NhGH11 unempfindlich gegenüber Metallionen und chemischen Reagenzien ist, welche typischerweise in assoziierten industriellen Produktionsprozessen vorhanden sind. Von Lignin abgeleitete Phenolverbindungen wie Kaffeesäure, Zimtsäure, Cumarsäure und Ferulasäure können jedoch die Katalyse beeinflussen, indem sie die Xylanaseaktivität von NhGH11 hemmen.

Daher unterstreichen diese Daten das Potenzial von NhGH11 für industrielle Abbauprozesse von Biomasse, bei denen durch Katalyse von Xylan-Biopolymeren nachhaltige Biokraftstoffe und biobasierte kommerzielle umweltfreundliche Materialien hergestellt werden können.

1. Introduction

Lignocellulose is a major part of the plant cell wall which is composed of 15-20% lignin, 25-30% hemicellulose and 40-50% cellulose¹. All of these components are arranged in the form of a three-dimensional network in the plant cell wall². Hemicellulose consists of a hetero-polysaccharide called xylan, consisting of β -1,4 linked xylose with various branches of sugars i.e. arabinose, xylose, galactose, or organic acids i.e. acetic acid, ferulic acid, glucuronic acid³. Xylan is the second most abundant polysaccharide after cellulose which covers 15-30% of the hardwoods and 7-10% of the softwood, constituting altogether 33% of total lignocellulosic biomass on earth^{4,5}. Additionally, the lignocellulosic biomass from rice straw, wheat bran, wheat straw, sugarcane bagasse, barley, woody residues, softwood herbaceous plants, paper pulps, manures and waste materials from forestry can be used as a sustainable source of energy⁶.

1.1 Importance of lignocellulose biomass

Lignocellulosic biomass varies from other energy sources and is used today as a sustainable resource for the production of bioethanol and other bio-based chemicals^{7,8}. However, in the past, nearly 80% of the total energy was produced utilizing fossil fuels and now the remaining resources of fossil energy are about to end, particularly crude oil⁹. Therefore, due to declining of oil reserves, the world will face a significant energy crisis in the future that needs to be resolved by introducing alternative sources of energy in time. Bioethanol has gained significant attention in the automotive sector because of its utility as an octane booster and even as a fuel. Brazil and USA produce ethanol from sugar cane and corn on a large scale. However, major bottlenecks in the process of bioethanol production are technical issues, which do not promote and allow cost-effective and efficient lignocellulosic bioethanol production¹⁰. An integrated and committed approach is needed to establish a commercial scale production of lignocellulosic bioethanol, also to support the establishment of a cleaner global climate and more energy-efficient world¹¹. The international energy agency estimates that lignocellulose will be able to provide 5-15% of the overall required energy by the year 2035¹².

Biofuels are categorized into four groups: first-generation biofuels are produced from edible plants containing energy in the form of simple sugars, oils and cellulose. In this context, there is an ongoing debate about the competition of first-generation biofuels with food production and security. Therefore, the second-generation biofuels are produced from

lignocellulosic biomass; non-food materials from agricultural fields, including wheat straw, sugar cane bagasse and forest residues. Today, several initiatives are under development to maximize and enhance the production of second-generation biofuels. It is important to mention that only the first and second-generation biofuels have been commercialized so far. Therefore, efforts are needed to accelerate the generation of more advanced biofuels from non-food raw materials (<https://easac.eu/programmes/energy/>). The third-generation biofuels are produced from algal biomass. The procedures to obtain third-generation biomass on an industrial scale are under development. Algal biomass offers several advantages in terms of biofuel processing methods that focus on non-food crops and organic waste materials, also wastewater effluents can be used for algal growth. In parallel, there are some disadvantages like low efficiency in terms of the lipids extraction generated from biomass, the possibility of environmental contamination due to growth in open ponds and high energy requirements because the entire production cycle require mixing, filtration, and centrifugation at each stage. Therefore, there is a need to optimize the production of algal biofuels and corresponding separation techniques to remove particular non-fuel waste components^{13,14}. Fourth-generation biofuels are obtained by bioengineering of genetically modifying algae or cyanobacteria that can support high yields of biofuels along with environmental benefits of capturing and storing carbon dioxide to reduce the greenhouse effect. The corresponding production process requires fewer steps and the genetically modified algae can utilize solar energy to transform carbon sources into biofuels, thereby the cost of the fermentation process can be avoided¹⁵. The procedures to obtain fourth-generation biofuels are under development so far, and they are expected to be a breakthrough in the field of advanced biofuels^{15,16}. Biofuels include bioethanol, biodiesel, biobutanol and biogas, however, to focus on the utilization of enzymes for bioethanol production, the steps of bioethanol from all four biofuel generations are summarized in Fig. 1.1.

1.2 Biofuels versus fossil fuels

Human activities are continuously influencing the climate resulting in increased pollution. The concentrations of greenhouse gases including CO₂, N₂O and CH₄ are increasing to a substantial level due to the modernization of human life and resulting in the greenhouse effect and global warming¹⁷. Additionally, utilizing conventional fossil fuels can enhance the toxicity of SO₂ gas, which causes respiratory problems and also contributes to acid rain. The shift from non-renewable fossil fuels to sustainable lignocellulose biofuels

ensures the reduction of pollution and global warming¹⁸. The concepts of biofuel production are based on principles of green chemistry that can enhance energy production from low-cost sources and recycling processes by utilizing the particular lignocellulose waste materials. Biofuels are fully biodegradable fuels that burn cleaner than fossil fuels, resulting in less emission of greenhouse gases; for example, even bioethanol has the potential to reduce greenhouse gases up to 86% that can consequently reduce environmental pollution (<https://www.energy.gov/sites/>). In this context, there is a need to identify efficient enzymatic cocktails, cheap substrates and cost-effective pre-treatment methods to optimize biofuel production, these optimizations will improve the production yields and utilization of side products that can lead to the minimization of the overall cost for successful commercialization of biofuels in future¹⁹.

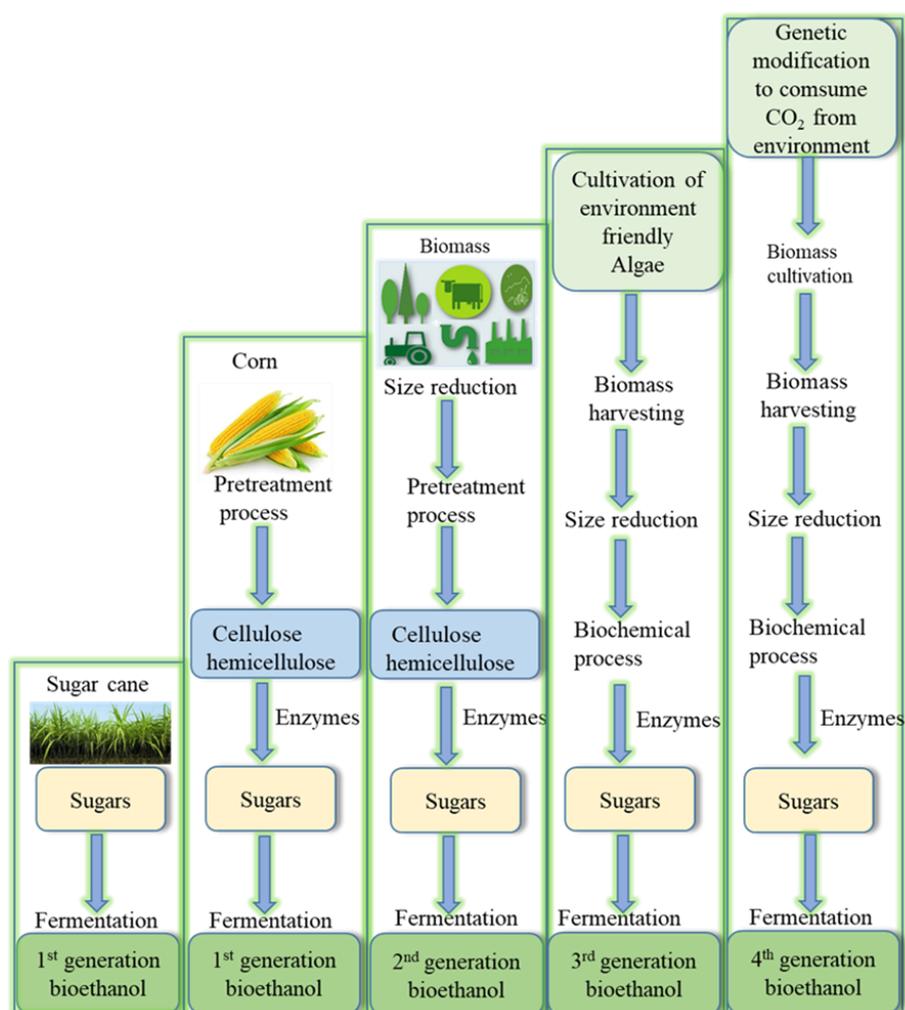


Fig. 1.1 Steps of bioethanol production. Biomass complexities increase the challenges in bioethanol production. Sugar cane is the easiest and most convenient raw material for bioethanol production. Corn or advanced biomass from different sources requires more steps for bioethanol production. Figure adapted from source, <https://slideplayer.com/slide/2562799/> with modification.

1.3 *Nectria haematococca*: a source of enzymes for biofuel production

The fungus *Nectria haematococca*, commonly called *Fusarium solani* (asexual name) is a filamentous fungus in the division of Ascomycota^{20,21}, genetically encodes 67 cellulases and 6 xylanases from GH10 and GH11 families (www.uniprot.org). *F. solani* consists of about 50 phylogenetic species commonly known as the *F. solani* species complex which are all saprophytic fungi found as pathogens to nearly a hundred genera of plants and animals, also causative agents of human mycoses, especially in immunocompromised patients²². *F. solani* is also associated with opportunistic fungal infections, especially eye infections in humans^{23,24}. Zhang and coworkers isolated *F. solani* species complex from the hospital environment, the isolates were found to be responsible for infections in humans and were designated in four different groups based on the phylogenetic relationship of their genome (Fig. 1.2)²³. *F. solani* has the largest fungal genomes with many repeated sequences that are enriched in unique genes. The studies on chromosomes suggest that several genes are involved in niche adaptation²⁵. Therefore, *F. solani* can grow in a variety of habitats i.e. agricultural regions, littoral zones, prairies, forests, deserts, scrub communities and coastal zones²³. Because of their diversity, some members of *F. solani* have been used for the biological control of plant parasites and pathogens²⁶.

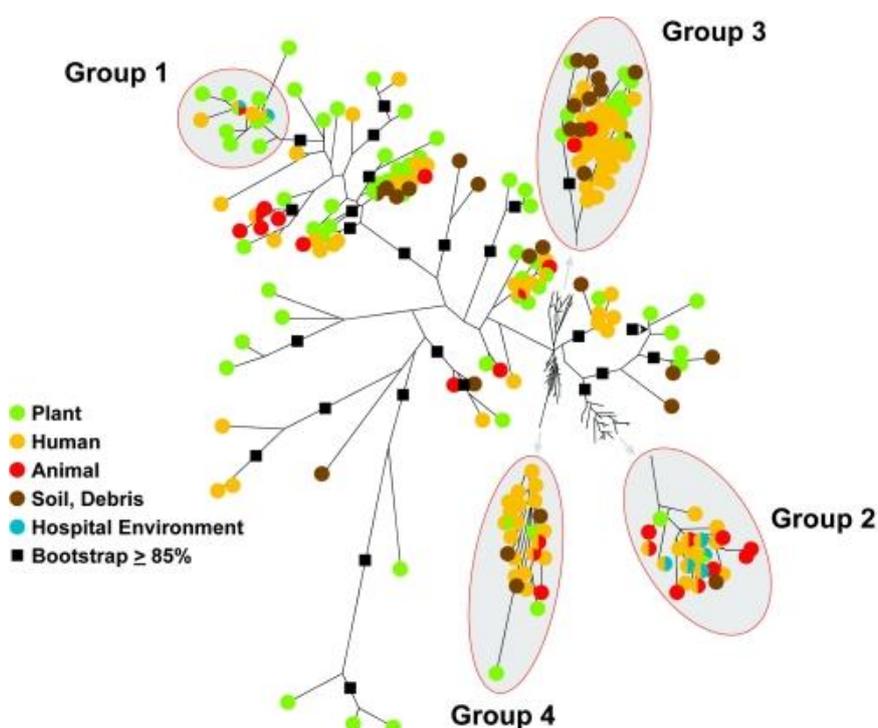


Fig. 1.2 The phylogenetic tree. Division of *F. Solani* isolates into designated groups 1 to 4 which indicate the phylogenetic diversity of fungus genome isolated from human infections, plant and animal sources, soil debris, and hospital environment. Figure source: doi:10.1128/JCM.00120-06.

1.4 Optimum conditions of carbohydrate degrading enzymes

Cost-effective degradation of plant biomass is an emerging need for a growing economy. Therefore, enzymatic cocktails need to be optimized for industrial processes. Considerable factors for the formation of industrial enzyme cocktails are shown in Fig. 1.3. The selection of suitable substrate, fermentation conditions and methods of pretreatment are the most important factors for an efficient breakdown of biomass material. Biochemical and structural studies of enzymes are providing knowledge to understand catalysis and optimization of catalysis parameters. Optimization of different carbohydrate-active enzyme ratios and the optimum conditions of each enzyme in a selected and designed cocktail should be known before commercial use²⁷. Also, optimization of thermostability is important because the enzymatic hydrolysis at a high temperature improves the efficacy of the overall process. Most of the GH11 xylanases are stable over a wide range of temperatures. GH11 xylanases characterized so far have optimum temperatures between 35 °C to 85 °C, categorizing them as mesophilic, thermophilic and hyper-thermophilic xylanases²⁸.

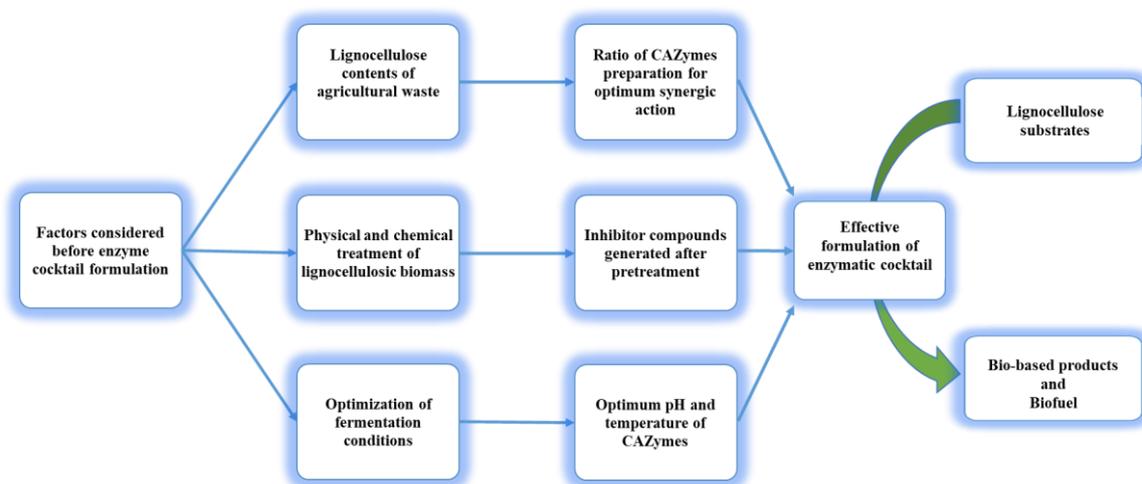


Fig. 1.3 Parameters for the efficient formulation of a biomass-degrading enzymatic cocktail. Selection of substrates, fermentation conditions and methods of pretreatment depend on the type of industrial application and need optimization at each step. Figure adapted from source: doi:10.1016/j.btre.2020.e00525.

Mesophilic xylanases have an optimum temperature of about 65 °C and include known xylanases from *Bispora* sp.²⁹, *C. thermocellum*³⁰, *P. varioti*³¹, *Streptomyces*³² and *T. lanuginosus*³³. Thermophilic xylanases have an optimum temperature between 70-80 °C and include xylanases from *Bacillus* sp. JB-99³⁴, *Caldicellulosiruptor* sp. Rt69 B.1³⁵, *C. stercorarium* XynA³⁶, *H. grisea* XYN2³⁷, *Paenibacillus* sp. DG-22³⁸ and *T. fusca* NTU22³⁹. Hyper-thermophilic xylanases have an optimum temperature of about 80 °C and

include xylanases from *T. flexuosa*⁴⁰ and *D. thermophilum*⁴¹. They have a longer half-life at higher temperatures than the thermophilic group. In addition to this, the use of high temperature in biorefineries also has benefits such as accelerated catalysis, reduced slurry viscosity and prevention of microbial contamination. Hydrolysis of lignocellulose biomass at elevated temperature accelerates the decomposition of xylan layers on top of the cellulose and increases the accessibility of cellulases⁴².

1.5 Industrial applications of xylanases

Approximately 20% of the world's enzyme market depends only on glycosyl hydrolyses, including xylanases, cellulases and pectinases⁴³. The commercial production of xylanases is ongoing in Germany, China, India, Denmark, Japan, Finland, Ireland, USA and Canada. The fungal genera e.g. *Fusarium*⁴⁴, *Penicillium*⁴⁵, *Aspergillus*⁴⁶, and *Trichoderma*⁴⁷ are used to produce xylanases. Some acidic xylanases producing fungal species, *Fusarium proliferatum*⁴⁴, *Penicillium oxalicum*⁴⁵, *Aspergillus niger*⁴⁶, *Trichoderma reesei*⁴⁷, *Paecilomyces thermophila*⁴⁸, *Hemicola insolens*⁴⁹, *Aspergillus oryzae*⁵⁰, *Pichia stipitis*⁵¹ and *Trichoderma virens*⁵² are being used for the production of xylanases. The comprehensive enzymes database named BRENDA provides the information of 190 entries of endo- β -1,4-xylanases with potential application in biofuel production, biotechnology, food and paper industry⁵³.

1.5.1 Food and feed technology

Glycoside hydrolases have many applications in the food industry. Among glycoside hydrolases, xylanases are mainly used in baking factories and are also being added to cereals to reduce the processing time and to improve the product quality. Xylanases from different families are used together to optimize the crumb structure of bread and to reduce the stickiness of the dough. They also improve bread loaf volume up to approx. 25% under specific conditions and enhance the shelf life of bread⁵⁴. Xylanases are also recommended to improve the texture of crackers and biscuits^{55,56}. They are used to improve the texture, taste and color of crispy wafer made by black gram by hydrolysis of arabinoxylan in the flour⁵⁷. They are also used in combination with amylases, cellulases and pectinases to improve stabilization, aroma, mineral salts and vitamins in fruit and vegetable juices⁴³. The synergic action of GH10 and GH11 xylanases induce the separation of wheat flour into starch and gluten⁵⁸. The quality of bread can be improved by gluten protein agglomeration via the addition of xylanases⁵⁹.

Xylanases have important applications in the animal food industry and poultry diet production, also supplementary xylanase cocktail is added to the animal diet to provide efficient growth effects which improve digestion, enhance immune parameters and support the gut microflora⁶⁰. A xylanase cocktail from *T. viride* is reported to enhance the digestibility of crude fiber by 60% and also enhancing the bodyweight of broiler chicken⁶¹. Selected commercially available xylanases are summarized in table 1.1, with source names and applications.

Table 1.1 Commercially available xylanases⁴³

Commercial Name	Organism	Distributors	Industrial Applications
Amano 90	<i>A. niger</i>	Amano Pharmaceutical	Pharmaceutical analysis, food industry
Allzyme PT	<i>A. niger</i>	Alltech	Feed additive for chickens and turkeys
Bio-Feed Plus	<i>H. insolens</i>	Novo Nordisk	Animal feed
Bleachzyme	Not cited	Biocon, India	Cellulose and paper industry
Cartazyme	<i>T. fusca</i>	Clariant, UK	Cellulose and paper industry
EcopulpX-200	<i>T. reesei</i>	Primalco	Cellulose pulp bleaching
Ecosane	Not cited	Biotec	Animal feed
ROHALASE®	Not cited	AbEnzymes	Baking, pasta, fruits and vegetables juices and wines
Solvay pentonase	<i>T. reesei</i>	Solvay Enzymes	Starch and bread-making
Sternzym HC 46	Not cited	Stern-Enzym	Bread preparations
Sumizyme X	<i>T. koningii</i>	Shin Nihon	Manufacture of mushroom and vegetable extracts, bread preparation, enzymatic peeling of cereals, animal feed
Xylanase	Not cited	Granotec	Weight decreasing in cream-crackers for better texture and taste
Xylanase	<i>Aspergillus</i>	Microbiosci	Sugar Industry, a feed additive for aquatic animals

1.5.2 Xylooligosaccharides

Xylooligosaccharides (XOSs) are degradation products of xylanases. The structure and yield of XOSs depend on the xylanase specificity applied for hydrolysis and biomass quality⁶². XOSs are novel prebiotics being used in the pharmaceutical, cosmetic, agriculture and food industries.

These prebiotic XOSs with different substitutions and degrees of polymerization can be obtained after xylanase hydrolytic action on biomass sources like wheat bran⁶³. Due to medical properties and health benefits, XOSs are also used as food supplements. XOSs are functional compounds having potential applications to control gastrointestinal disorders, diabetes, obesity and intestinal cancer⁶⁴. These health benefits of XOSs are monitored in rats after enrichment of the bifidobacteria level in the intestine and feces⁶⁵.

1.5.3 Paper industry

Xylanases are also important for the paper industry, they hydrolyze the linear 1,4- β -polyxylose linkages between cellulose and lignin. The hydrolysis of xylan linkages makes the paper pulp structure more permeable for other enzymes. The action of xylanases allows the better extraction of lignin that may also enhance the bleachability of paper and can decrease the amount of total bleaching chemicals (chlorine and chlorine dioxide) applied in the paper pulp industry⁶⁶. Xylanases can improve the final brightness of the paper⁶⁷. In this context, thermostable xylanases combined with endoglucanases are used for biobleaching to improve paper quality^{68,69}.

1.5.4 Cleaning agents

Proteases, cellulases, lipases and amylases are eco-friendly and effective anti-felting cleaning agents for wool as compared to conventional soaps. Also, xylanases from various sources and families are proved as important substitutions of conventional chemical soaps⁷⁰. For example, xylanases along with pectinases can be used to clean the fibers as efficiently as soap, also these enzymes can replace soaps due to their inertness towards physical damage of wool⁷¹.

1.5.5 Wine and juice industries

The enzymatic cocktails containing cellulases, xylanases and pectinases are used in wine and juice industries because of their ability to hydrolyze small unwanted colloidal particles and pectins present in the fruit extracts. Several commercial enzymatic cocktails are

widely used in wine production due to the benefits of maximizing juice yield, aroma, flavor and color of wines⁷²(Fig. 1.4). Xylanase from *S. sclerotiorum* is well known for juice clarification by hydrolysis of oligosaccharides into monosaccharides and by decreasing the level of insoluble materials up to 27 percent⁷³. The enzymatic cocktails are mostly produced by fungi, e.g. primarily from *Aspergillus*, *Rhizopus* and *Trichoderma*, to be applied in the wine industry to improve wine quality and stability⁷⁴.

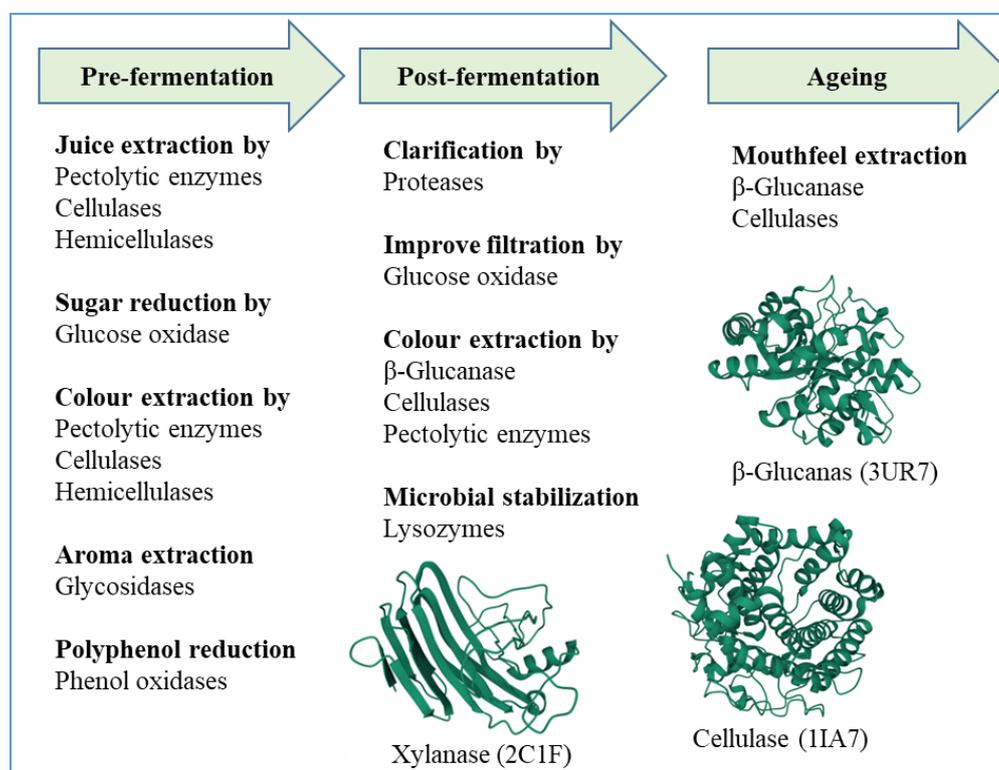


Fig. 1.4 An overview of enzymes used in wine production and their applications⁷². Commercial enzyme cocktails are widely used in the pre-fermentation, post-fermentation and ageing process of wine. These enzymes enhance the quality of wine. Figure modified from <https://doi.org/10.1007/s13197-020-04489-0>.

1.5.6 Biofuel

Bioethanol (C_2H_5OH) is a worldwide well-accepted biofuel that can be produced from lignocellulose biomass. In the production process of lignocellulosic bioethanol, the lignocellulose biomass is subjected to four major steps which are pretreatment of lignocellulose biomass, enzymatic hydrolysis of complex lignocellulose biomass, fermentation process and separation of bioethanol⁷⁵.

Pretreatment: involves the mechanical and/or chemical breaking of the complex structures in lignocellulose biomass to reduce the size and complexity.

Pretreatment is important to reduce the degree of polymerization and crystallinity of cellulose, therefore, it increases the exposed surface area of raw material to make them suitable for enzymatic hydrolysis⁷⁶. Pretreatment of biomass can be performed by a single or combination of the following methods:

Mechanical pretreatment: can be done by milling to reduce particle size and crystallinity of lignocellulose biomass. The particle dimensions of biomass can be reduced to approx. 0.2 mm. Different types of motorized equipment are used for size reduction by milling methods. The final particle size and crystallinity of biomass depend on the type of milling method and the chemical nature of biomass. Pretreatment by milling does not produce toxic materials or phenolic compounds which are potent inhibitors of lignocellulose degrading enzymes. Mechanical pretreatment can be done via microwaves and has several advantages like easy operation and they are economically favorable⁷⁷. Mechanical pretreatment can also be performed by hot water steam. This method enhances the rates of enzyme hydrolysis by maintaining the pH between 4 to 7 and reduces the formation of inhibitors i.e. aldehydes and phenolic compounds from hemicellulose⁷⁸.

Chemical pretreatment: involves utilizing strong/dilute acids, alkaline solutions, ionic liquids, oxidizing agents, or organic solvents to reduce the complexity of lignocellulose biomass. Chemical pretreatment has limitations including toxicity for cellulases and xylanases, this process also requires expensive chemicals and special corrosion-resistant equipment. However, also has the advantage of minimizing heat and energy requirements⁷⁹.

Biological pretreatment: is an environmentally friendly and cost-effective procedure. In this method, biomass is treated with lignin-degrading fungi or bacteria by applying entire cells or only enzymes for degradation of lignin⁸⁰. Fungi are most suitable to be applied for this process, as they can produce a complete cocktail of cellulose, hemicellulose and lignin-degrading enzymes⁸¹.

Enzymatic hydrolysis: involves the depolymerization of the insoluble cellulose and hemicellulose into a mixture of hexose and pentose sugars, which are subsequently fermented into bioethanol. The enzymatic cocktail contains cellulases, hemicellulases and lytic polysaccharide monooxygenases (LPMOs) to catalyze the hydrolysis of biomass (Fig. 1.5). Cellulose degrading enzymes include endo- β -glucanases, exoglucanases, cellobiases and β -glucosidases^{82,83}. Fungal cellulases contain exoglucanases including β -1,4 glucosidase β -1,4 glucan cellobiohydrolases and β -glucosidases that hydrolyze 1-4-

linkages at the ends of cellulose. While endoglucanases and LPMOs randomly cleave the crystalline cellulose and produce free chain ends which are hydrolyzed by exoglucanases⁸⁴. Furthermore, exoglucanases cleave the polysaccharides into cellobiose subunits which are hydrolyzed by cellobiohydrolases⁸⁵.

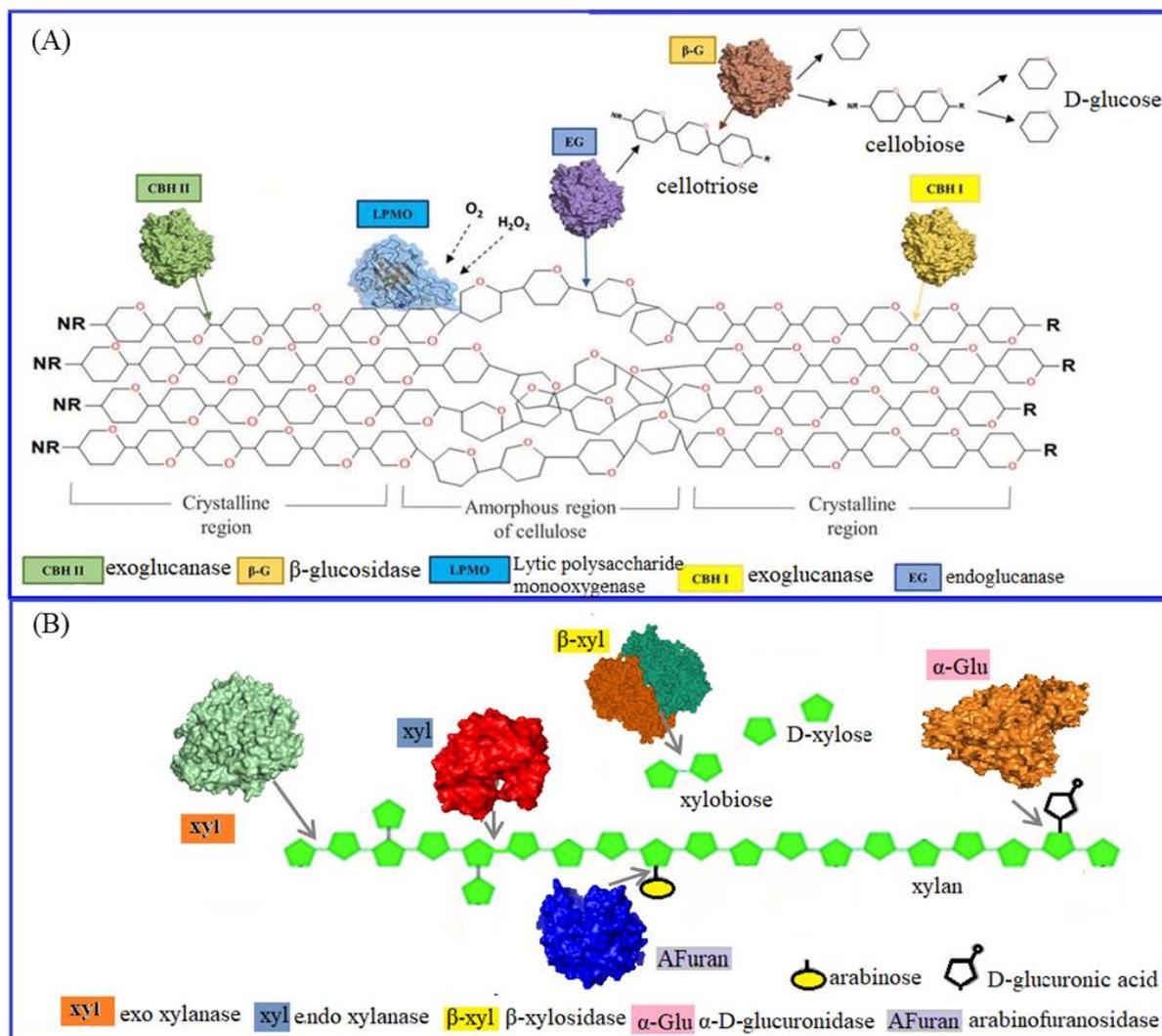


Fig. 1.5 Enzymatic hydrolysis of lignocellulose biomass. (A) Structure of crystalline and amorphous regions of cellulose with cellulose-degrading enzymes. (B) Structures of hemicellulose (xylan) with indicated positions for enzyme action at different sites. Image modified from the source: <https://doi.org/10.1002/elsc.201800039>.

Hemicellulose degrading enzymes can be divided into two categories: enzymes having endo-activity acting in the middle of polymers and enzymes having exo-activity acting on ends of the xylan polymers. However, many hemicellulases own both kinds of activities. Fungal hemicellulases include β -1,4 xylanases, β -1,3 xylanases, β -xylan glucosidases and β -mannosidases.

Besides these debranching enzymes can act on variously branched xylan polysaccharides⁸⁶. Debranching enzymes are a kind of auxiliary enzyme that can also act on hemicellulose. Debranching enzymes can be divided into enzymes acting on ester bonds i.e. α -l arabinofuranosidases, α -glucuronidases, and enzymes acting on glycoside bonds i.e. β -xylosidases, acetyl-xylan esterases, ferulic acid esterases and coumaric esterases^{87,88}.

The commercial cocktails of enzymes for biofuel production additionally comprise endoglucanases and cellobiohydrolases, isolated from different aerobic fungal stains⁸⁹. Therefore, with the synergistic action of all these enzymes, the components of lignocellulosic biomass can be hydrolyzed to monosaccharides, which can be further used for biofuel production.

The carbohydrate-active enzymes (CAZy) is a special online database that was oriented in the year 1998 to provide information about enzyme families that can hydrolyze, modify, or create glycosidic bonds. Many families of CAZy are involved in the hydrolysis of lignocellulose biomass, as summarized in table 1.2 with their preferable substrates. These families of enzymes can be utilized in enzymatic cocktails used for biofuel production.

Table 1.2 Carbohydrate-Active enzyme (CAZy) for biofuel production

Substrate	Enzyme	CAZy families	Function
Cellulose	Endo- β -1,4-glucanase	GH5, 6, 7, 8, 9, 12, 44, 45, 48, 51, 74, 124, 131	Catalyze the cleavage of internal bonds randomly at amorphous regions of cellulose
	Cellobiohydrolase (CBHI and II)	GH6, 7, 9, 48	CBHI acts at the reducing and CBHII at the non-reducing chain ends of cellulose
	β -1,4-glucosidase	GH1, 3	Hydrolysis of oligosaccharide at the non-reducing chain ends of cellulose to release β -D-glucose
	LPMO	AA9, AA10	Catalyze the oxidative cleavage of cellulose
	LPMO	AA13	Catalyze the oxidative

			cleavage of starch
Xylan	endo-1,3 and endo-1,4- β -xyylanase	GH8, 10, 11, 8	Hydrolysis of endo- β -1,4-xylosidic bonds
	Glucuronoarabinoxylan endo- β -1,4-xyylanase	GH30	Hydrolysis of endo- β -1,4-xylosidic links in glucuronoarabinoxylans
Xyloglucan	endo- β -1,4-glucanase/ α -arabinofuranosidase	GH5, 9, 12, 16, 44, 74, 43, 51, 54, 62	Hydrolysis of endo- β -1,4 glucosidic bonds in xyloglucan.
	α -galactosidase	GH27, 36	Hydrolysis of α -D-galactose residues at non-reducing ends of xylan
	β -1,4-galactosidase	GH2, 35, 42	Hydrolysis of terminal non-reducing β -galactose residues
Starch	α -amylase	GH13, 57	Hydrolyze the internal α -1,4-glycosidic bonds in amylose and amylopectin
	Glucoamylase	GH15	Act on non-reducing ends of starch to hydrolyze both, α -1,4 and α -1,6 bonds in amylose and amylopectin
	α -1,4-glucosidase	GH31	Act on non-reducing ends of starch to release α - glucose
	β -amylase	GH14	Act on non-reducing ends of starch to hydrolyze α -1,4-glycosidic bonds to remove maltose units

Fermentation: the monosaccharides released by enzymatic hydrolysis of biomass are taken up by microorganisms for the fermentation process. Microbial fermentation of sugars is carried out by bacteria, fungi, or yeasts, and the products of the fermentation process depend on the selected microorganism. *S. cerevisiae* and *Z. mobilis* are commonly used microorganisms for bioethanol production but they can only utilize hexose sugar under standard conditions⁹⁰. However, *S. cerevisiae* can utilize pentose sugars for ethanol

production but only under special growth conditions⁹¹. *S. cerevisiae* are used at commercial levels due to approx. 90% bioethanol yield, high ethanol tolerance, simple and inexpensive growth media and wide pH range. Also, they are resistant to several inhibitors and contaminants present in growth media⁹².

Moreover, the facultative fermentative yeasts like *P. stipitus*, *K. marxianus*, *P. tannophilus* and *C. shehatae* are suitable for fermentation of pentose sugars, but also with certain limiting factors⁹³. The edible *Ascomycetes* fungus can also be used as a potential fermentation candidate, due to its xylose-fermenting capability⁹⁴. Bioethanol can be produced by three different fermentation processes e.g., batch, fed-batch, or continuous process⁹⁵.

In process of batch fermentation, at the beginning of the whole batch all growth ingredients, feedstock, microorganisms and nutrients are added to the fermentation vessel followed by bioethanol recovery. This fermentation process does not require skilled workers due to fewer control requirements, easy management of feedstocks and sterilization process as well as low cost is required. Due to these advantages, batch fermentation is commonly used in bioethanol research^{96,97}.

However, in the fed-batch process, ingredients are added stepwise during the fermentation process, this process is employed in several industrial productions because of the dual benefits from batch and continuous fermentation processes. Industrial adaptation of this process is due to the following advantages such as maximum lifespan and cell viability, high product accumulation, less inhibition by a high concentration of substrate and proper control of physical factors like pH, temperature and oxygen level⁹⁸.

The continuous fermentation process needs constant input of growth ingredients and removal of products from the vessel to maintain the ongoing fermentation process⁹⁹. The process can be carried out in two basic types of reactors; continuous stirred reactor and plug flow reactor. This process has advantages over batch fermentation due to saving time as refilling of ingredients and withdrawal of products is possible without vessel cleaning that increases productivity and decreases the overall cost¹⁰⁰.

Separation process of bioethanol: bioethanol obtained by fermentation microorganisms must be purified for further applications. Bioethanol separation is very crucial but at the same time, it is very expensive due to high energy requirements. Several separation processes like extractive distillation, pervaporation and conventional molecular sieves can

be applied to obtain a better quality of bioethanol¹⁰¹. However, distillation is a costly procedure to obtain an azeotropic mixture of ethanol (96.6%) and water (3.5%). Today there are some conventional procedures known for the separation of bioethanol, which includes the combination of liquid-liquid extraction and extractive distillation¹⁰². Moreover, some alternative configurations of ethanol purification in presence of thermal coupling have low energy consumption and are environmentally friendly in terms of greenhouse effect¹⁰³.

1.6 Introductory preface of X-ray crystallography

The German physicist Wilhelm Conrad Röntgen discovered X-rays in 1895, while testing the cathode rays he noticed unknown rays coming from a chemically coated tube.

He named them X-rays because of their unknown nature¹⁰⁴. Later on, he obtained the Nobel Prize in 1901 due to his revolutionary discovery. After the discovery of X-rays, they are being used for imaging of bones and cavities in soft tissues¹⁰⁵. However, the application of X-ray in the field of crystallography was started thirteen years later when Max von Laue observed X-ray diffraction by copper sulfate crystals¹⁰⁶, now known as Laue-crystallography in honor of Max von Laue¹⁰⁷. Although, Sir William Lawrence Bragg with his father in 1915, provided the interpretation of Laue's diffraction as reflections of a discrete lattice plane from sodium chloride crystal and formulated the fundamental law of diffraction, known as Bragg's Law:

$$n\lambda = 2d_{hkl} \sin \theta$$

Where λ is the wavelength of the x-ray, d is the spacing of the crystal layers (path difference), θ is the incident angle (the angle between the incident ray and the scattering plane), and n is an integer.

Pioneers of protein crystallography were Max Perutz and John Kendrew, who obtained the Nobel Prize in 1962 for solving the first X-ray crystal structure of myoglobin¹⁰⁸ and hemoglobin¹⁰⁹, after that the protein crystallography started to develop further. Afterwards, the synchrotron radiation sources were developed for producing the most intense X-rays. Synchrotron facilities and beamlines were built around the machines that were designed for particle physics studies, also included an EMBL outstation at DESY in Hamburg. A synchrotron storage ring was constructed with experimental sectors, shown in Fig. 1.6.

In the 1960s, synchrotron radiation was considered a high-cost by-product from particle accelerators. Later on, synchrotron radiation was found to be completely favorable for X-ray crystallography studies¹¹⁰.

The first generation synchrotron was built in the 1970s followed by the second generation in the 1980s, currently, the third generation is also near to replace by the upcoming fourth-generation synchrotrons (<https://www.inls.cnpem.br/the-inls/history-of-the-synchrotron>). High flux synchrotron radiation sources decreased the required dimensions of crystals to measure the reliable Bragg intensities¹¹¹ and improved resolutions.

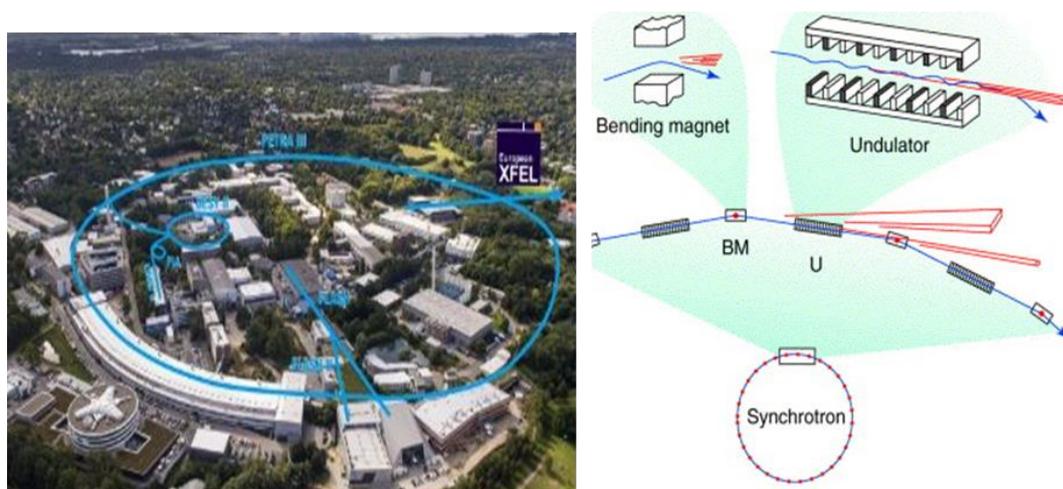


Fig. 1.6 Deutsches Elektronen Synchrotron (DESY). DESY is an international center of fundamental research founded in 1959. The right figure is indicating a scheme of how synchrotron radiation is produced, via bending magnet or undulator beamline¹¹² (Image Source: <https://www.research-in-germany.org>). Electrons and/or positrons which are accelerated and traveling close to the speed of light, kept on a track via a strong magnetic field instantaneously emit X-rays tangential to the horizontal plane they travel¹¹³.

Since 1971, the Protein Data Bank (PDB) is serving as the single repository of information about the three dimensional structures of nucleic acids, proteins and complex assemblies. Over the past few decades, the PDB is indicating an exponential growth of X-ray structures (<https://www.rcsb.org/stats/growth/growthx-ray>) and the numbers of X-ray structures almost exponentially increased from 13 to 149383 from 1976 up to 2021. The PDB contains today approx. 90% of structures are solved by using synchrotron radiation and the numbers of deposited structures are increasing every year, as shown in Fig. 1.7.

High-resolution structures of macromolecules are important to understand the mechanisms of biological functions. The advancements of synchrotron beamlines and modern techniques of molecular biology improved the crystallographic studies of macromolecules.

Synchrotrons have remarkable features of providing highly brilliant X-ray beams at a wide range of wavelengths¹¹⁴.

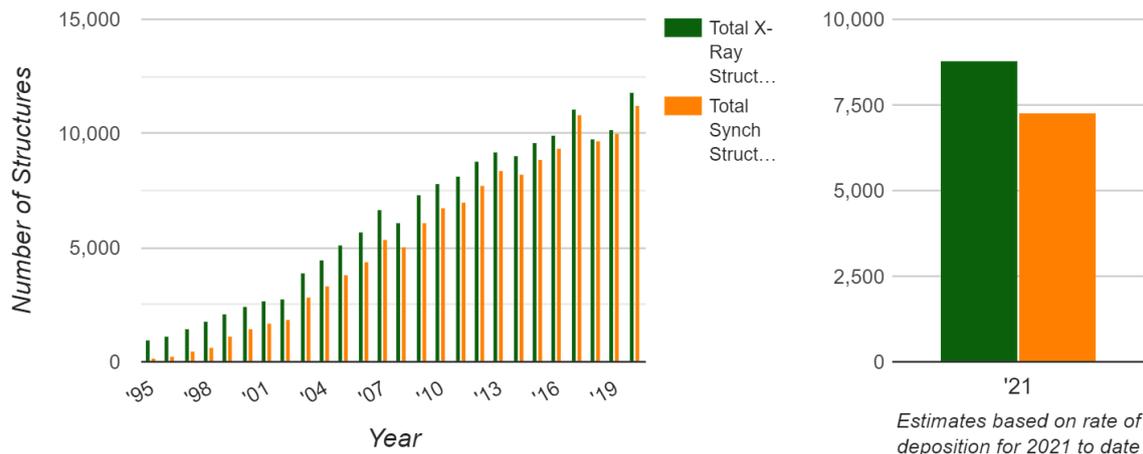


Fig. 1.7 Protein Data Bank structure depositions determined by X-ray crystallography. Total vs. structures solved applying data collected with synchrotron radiation (<http://biosync.sbkb.org/>).

1.6.1 Conventional vs serial crystallography

Conventional crystallography is the most productive method in structural biology where a suitable single crystal of a protein is irradiated with X-ray¹¹⁵. Single crystal diffraction data collection requires crystals with approx. dimensions between 50 to 500 μm , to be mounted on a goniometer and recording diffraction patterns at cryogenic temperature (Fig. 1.8 A). Single crystal diffracts X-rays according to Bragg law into discrete diffraction spots, also called reflections. The diffraction data provides the amplitudes of the Fourier transform of the electron density in the unit cell and the missing phases can be determined by phasing methods e.g., isomorphous replacement methods i.e. single isomorphous replacement (SIR) or multiple isomorphous replacement (MIR), anomalous diffraction methods i.e. single-wave anomalous diffraction (SAD), multi-wavelength anomalous dispersion (MAD) and molecular replacement (MR) method¹¹⁶. In SIR or MIR, the diffraction data are collected from the native crystal and several derivatives crystals by soaking in a heavy atom solution. The scattering differences of additional heavy atoms present in the derivative crystals, relative to the native protein are used for phase information. However, SAD and MAD techniques utilize one or more data sets, recorded from only one crystal containing suitable anomalous scatters. Thereby, the protein phases are calculated from the wavelength-dependent quantitative differences in the anomalous

scattering contribution of certain atoms contained in the crystal¹¹⁷. In the molecular replacement method, the phases from the similar structure with similar folds are used to substitute the missing phase information. Molecular replacement is the most commonly used method to solve phase problem in X-ray crystallography¹¹⁸. Conventional crystallography of macromolecules provided the majority of macromolecular 3D structures available to date (<https://www.rcsb.org/>).

Along with the development of new radiation sources, a revolutionary serial femtosecond X-ray protein crystallography method has been developed for macromolecular structure determination where X-ray diffraction data are collected by applying a stream of nanocrystals. Crystallographic studies at third-generation synchrotrons and free-electron lasers utilizing the serial crystallography approach extended the range of usable crystal sizes to approximately 10 μm , also providing the benefits of less radiation damage and room temperature data collection for ligand binding studies¹¹⁹. The dynamic studies of protein-ligand interactions at high resolution were the major challenges of X-ray crystallography, now in principle possible by serial crystallography methods. As an example, the experimental setup of a tape drive device for mix-and-diffuse serial crystallography experiments is shown in Fig. 1.8 B.

Electron densities of macromolecules complexed with ligands can be calculated using diffraction data obtained at mixing times of milliseconds, allowing for example to follow the catalytic mechanism of macromolecules¹²⁰. Applying the tape drive method, the diffraction pattern collected from individual microcrystals can be merged into data sets that can further be used to analyze a macromolecular structure complexed with ligands^{119,121}. The tape drive approach of serial crystallography enables room temperature diffraction data collection from the suspension of small crystals¹²².

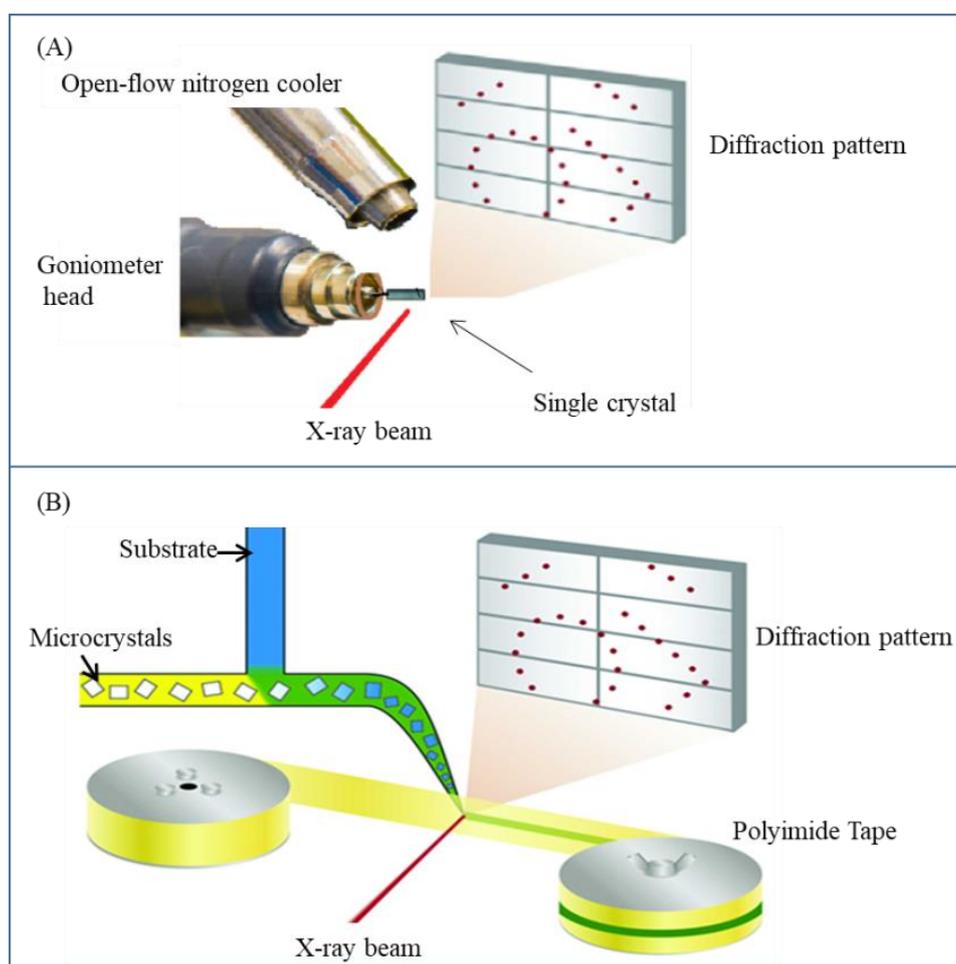


Fig. 1.8 Conventional vs serial crystallography. (A) A single crystal with approx. the dimension of $100\ \mu\text{m}$ is mounted on a goniometer for data collection under cryogenic conditions and the detector at a suitable distance is recording the diffraction data. (B) A micro-crystal suspension (yellow) mixed with a substrate solution (blue). The tape is fixed between two rollers under tension; the needle for sample delivery is fixed so that the X-ray beam passes micro-sized crystals¹²⁰, (B) adapted from <https://doi.org/10.1107/S2052252517013124>.

1.6.2 Crystallographic studies of xylanases

Xylanases can convert the complex form of xylan into simpler subunits. The first xylanase from *A. foetidus* was produced and purified in the year 1955 for biochemical analysis¹²³. Afterwards, many other xylanases isolated from different bacteria, fungi and yeasts were purified and characterized biochemically. After developments in structural biology, the first three dimensional structure of xylanase was solved in 1994^{124,125} an evolution of deposited 3D structures of xylanases started afterwards as shown in Fig. 1.9.

Xylanases are highly diverse enzymes at structural, biochemical and catalytic levels, additionally; a single organism can produce many different types of xylanases. For

example, fungi like *S. dimorphosporum* and *A. niger* can release approx. 15 different forms of xylanase respectively^{126,127}. The diversity of these enzymes originates from well-identified differences, by genetic redundancies, post-translational modifications and presence or absence of a signal peptide. The versatility provides a better adaptation to microorganisms by optimizing the biodegradation of plant cell walls¹²⁸.

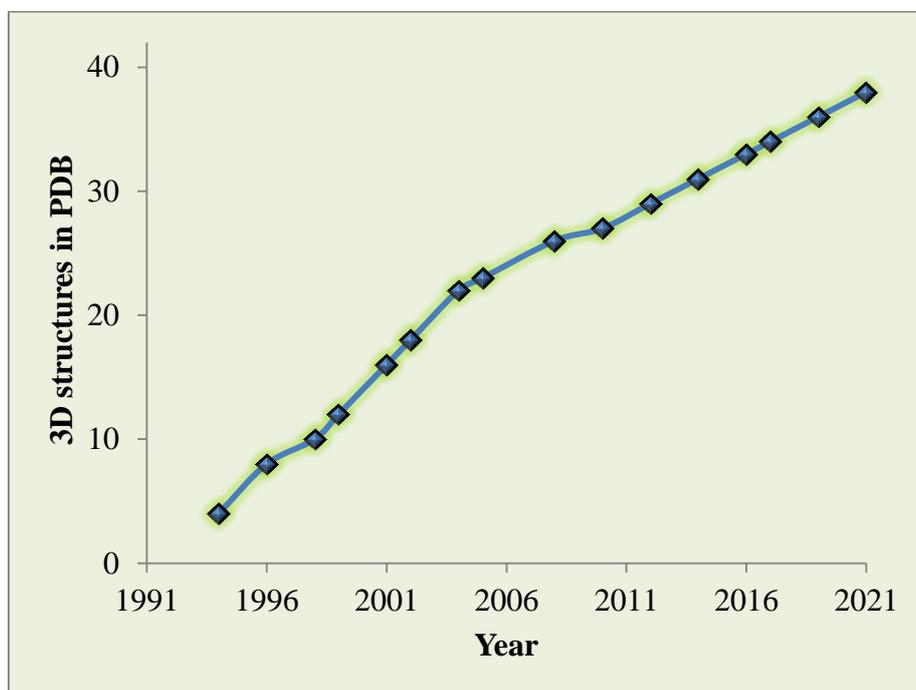


Fig. 1.9 Evolution of deposited PDB structures of GH11 xylanases (1994-2021). Data taken from the carbohydrate active enzyme database (www.cazy.org/GH11structure) under the description of family GH11 and dates of structure depositions are confirmed from the protein data bank (<https://www.rcsb.org>).

Glycosyl hydrolases (GHs) with xylanase activities are divided into six families GH5, GH8, GH10, GH11 and GH43. These families differ in structures, substrate specificities and mode of actions. GHs are further divided into clans of related families due to folding similarities in the 3D structures. For example, β -jelly roll xylanases are classified as clan GH-B and GH-C, while the xylanases with $(\beta/\alpha)_8$, $(\alpha/\alpha)_6$ and 5-blade β -propeller folds are classified as clans GH-A, GH-M and GH-F respectively. GH5 and GH10 xylanases belong to clan GH-A, while GH8, GH11 and GH43 belong to GH-B, GH-C, GH-M and GH-F respectively, (<http://www.cazy.org/>).

Catalytic mechanism: xylanases do not hydrolyze randomly but they decorate the substrate backbone selected for hydrolysis depending on the nature of the substrate¹²⁹. The families GH5, GH10 and GH11 consisted of xylanases, which have catalytic glutamate residues in the active site to hydrolyze the substrate with retention of the anomeric center

of substrate (<http://www.cazy.org/>). While GH8 and 43 have glutamate and aspartate as catalytic residues to hydrolyze with inversion of the anomeric center of substrate¹³⁰. The proposed catalytic mechanism of GH11 xylanases is shown in Fig. 1.10.

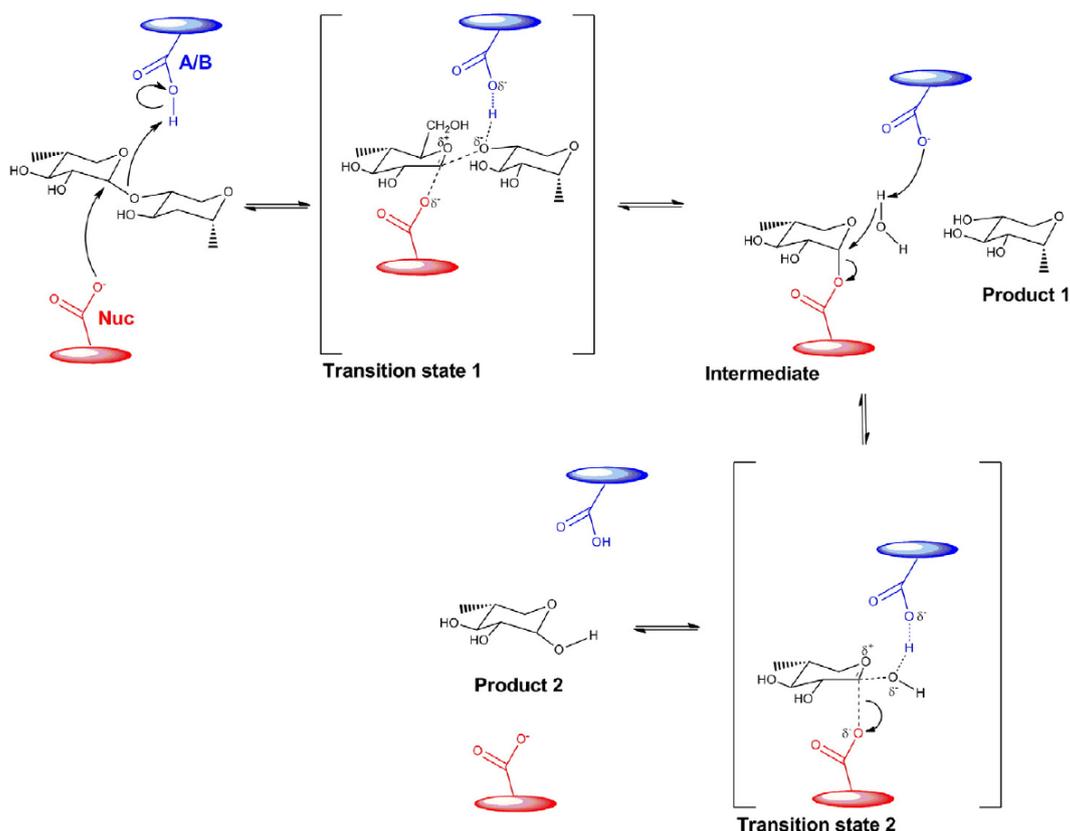


Fig. 1.10 Proposed catalytic mechanism of GH11 xylanases. In the first step, E80 acts as the nucleophile and attacks the anomeric center of glycosidic bonds in sugar polymers to form a glycosylated enzyme intermediate. In the second step, the deprotonated acidic carboxylate of E180 acts as a base and assists a water molecule to perform a nucleophilic attack to hydrolyze the glycosyl enzyme intermediate and to release the hydrolyzed product i.e. xylose sugar. Source: <https://doi.org/10.1016/j.biotechadv.2011.10.003>.

Substrate preferences: GH5 enzymes can catalyze the hydrolysis of glucuronoxylan or acetylated glucuronoxylan substrates¹³¹, while GH8 enzymes prefer a xylan with linear and mixed linkage (β -1,3 and β -1,4)¹³². GH11 xylanases are specific for hydrolysis of β -1,4 linked xylan substrates⁴. While GH10 xylanases show catalytic versatility for substrates^{128,133}. The substrate selectivity and specificity of xylanases are determined by the structural features of their active sites e.g. GH10 enzymes have four to five subsites⁴ but the substrate-binding clefts of GH10 xylanases have a variable number of subsites¹³⁴. GH11 and GH8 xylanases have an extended substrate-binding cleft with six subsites to accommodate six-xylose units in the center of the catalytic site^{132,135}.

1.7 Inhibition of xylanases

Due to pathogenic fungi and bacteria, plants show a defense response including the production of xylanase inhibitor proteins. These inhibitor proteins consist of two major classes of plant proteins named “Xylanase inhibitor proteins” (XIP) and “*Triticum aestivum* xylanase inhibitors” (TAXI)^{136,137}. These inhibitors are naturally produced by wheat, barley and rye as a part of their defense against fungal xylanases. These natural inhibitors can inhibit the xylanases when they are used in the food industries or for the hydrolysis of lignocellulose biomass. The inhibition of xylanase takes place when the XIP-loop binds between the active site and thumb region of GH11 xylanases¹³⁸ while a histidine residue of TAXI binds to the catalytic residues of GH11 xylanases to inhibit their activity¹³⁹. The bindings of XIP and TAXI with GH11 xylanases are shown in Fig. 1.11. TAXI has two forms: a non-cleavable form (TAXI-I) that consists of a 40 kDa polypeptide chain and a cleavable form (TAXI-II) with 10 kDa and 30 kDa polypeptide units. These different forms differ in their specificities and N-terminus sequences¹⁴⁰. TAXI-I inhibits GH11 xylanases with low and high isoelectric points, while TAXI-II only inhibits xylanases with high isoelectric points¹⁴¹.

In addition to protein inhibitors, lignin-derived phenolic compounds also inhibit several glycosyl hydrolases including xylanases. The pretreatment of lignocellulose biomass can result in the production of many lignin-derived phenolic acids that decrease the efficiency of biofuel production by inhibition of many important GH enzymes¹⁴². The most common pretreatment by-products are simple phenolic compounds, derivatives of phenol and furans^{143,144}. Besides these several other inhibitors like furan aldehydes e.g. furfural and 5-hydroxymethylfurfural or some aliphatic acids like formic acid, acetic acid, and levulinic acids generated from pretreatment or degradation of lignin can inhibit the enzymes used for biofuel production¹⁴⁵. Xylan from hemicellulose contains these phenolic compounds as branches on xylose sugars of hemicellulose, shown with yellow color in Fig. 1.11 C. Soluble phenolic inhibitors can also inhibit the fermentation process of sugars to bioethanol¹⁴⁶. Lignin degradation products, phenol derivatives, hydroxynamic acid and its derivatives, tannins and gallic acid are also potent inhibitors of many glycosyl hydrolases¹⁴³. Some other phenolic compounds like gallic acid, hydroxyl cinnamic acid, 4-hydroxybenzoic acids, vanillin and tannic acid caused 20 to 80% inhibition of glycosyl hydrolases¹⁴⁷ and the removal of these lignin-derived phenolic compounds from biomass can result in increased productivity of bioethanol¹⁴⁸. Lignin derived phenolic compounds i.e., coumaric acid, ferulic acid, 3,4-dimethoxycinnamic acid, 3,4,5 trimethoxy cinnamic

acid, 3-(4-hydroxyphenyl) propionic acid and ethyl 4-hydroxy-3-methoxy cinnamate also showed an inhibitory effect on glucosidase activity and the inhibition depends on the structure of lignin-derived phenolic compounds and their functional groups. Inhibition of glucosidase was not observed when phenolic free lignin was used for inhibition studies¹⁴⁹.

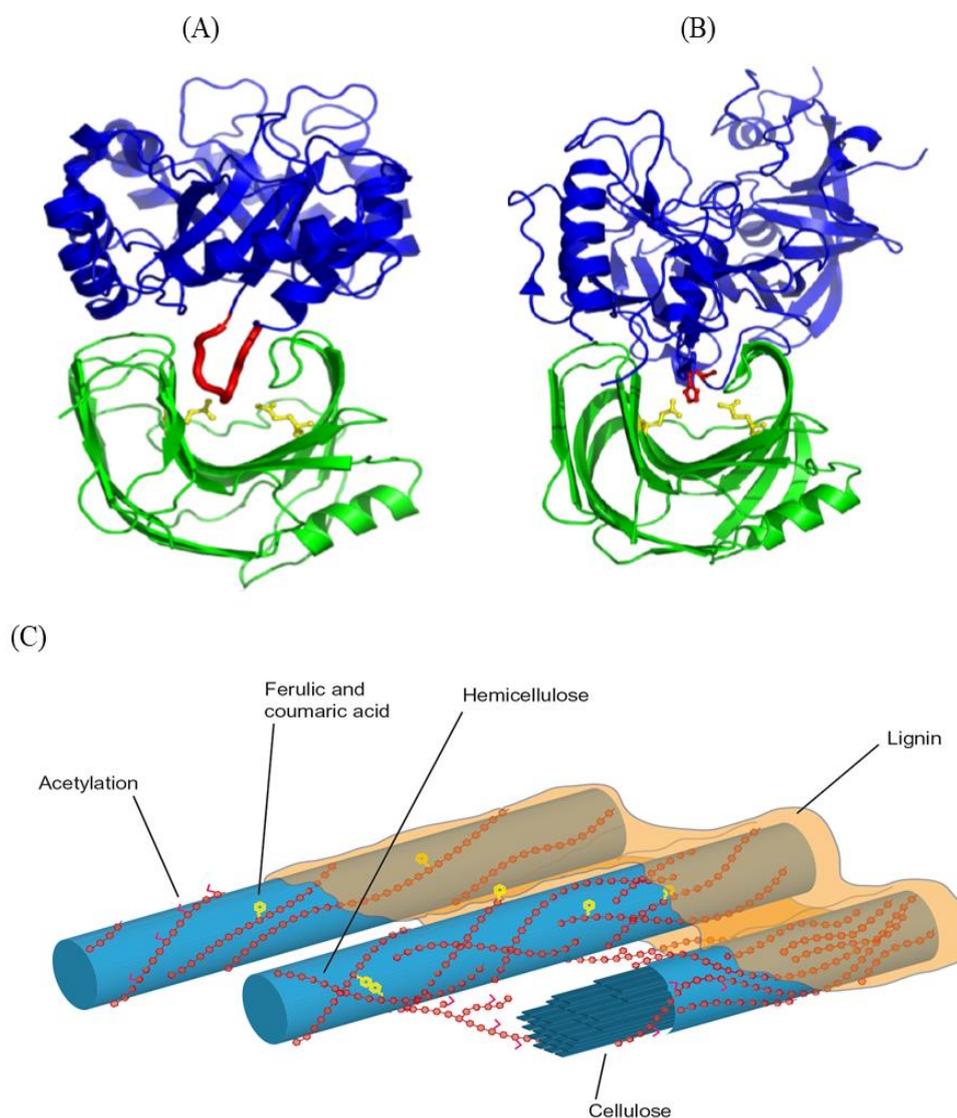


Fig. 1.11 Inhibitors of GH11 enzymes. (A) Structure of the inhibition complex of XIP and a GH11 xylanase¹⁵⁰. A loop (red color) from XIP binds between the active site region of GH11 xylanases (green) and the thumb region (B) Structure of the complex of TAXI and a GH11¹³⁹. A histidine from TAXI binds to the catalytic residues of GH11 xylanases (yellow). (C) Plant cell wall with structural components, cellulose and hemicellulose. Hemicellulose is variously branched with phenolic compounds (inhibitors) shown in yellow color. These phenolic compounds can inhibit GH11 enzymes¹².

1.8 Scope of the study

Agricultural and forest waste biomasses can be hydrolyzed by carbohydrate degrading enzymes to make them applicable for biofuel production. In this context, corresponding research activities have focused on analyzing the potential of lignocellulolytic enzymes able to convert waste biomasses into small carbohydrate molecules useful for bioethanol production. Xylanases are not only applied on enzymatic cocktails used in biofuel industries, but they are also equally important for food, pharmaceutical and paper pulp industries. In 2017, *Nectria haematococca* was identified and characterized by the group of Prof. Dr. Wilhelm Schäfer at the Department of Botany, University of Hamburg. After extensive bioinformatic studies, GH11 enzymes from *N. haematococca* were found to be homologous to xylanases already in use for industrial applications, also the carbohydrate active enzymes database (CAZy) revealed several putative xylanases from *N. haematococca* genome, which were structurally uncharacterized. Therefore, in this study xylanase from *N. haematococa* (NhGH11) was analyzed and characterized in details, starting from protein expression optimization, purification to complete biochemical and structural analysis.

The main objectives of this study can be summarized as:

- Selection of an expression vector, strains of bacterial cells, expression media and conditions for protein expression
- Optimization of solubility buffers, purification protocols and cleavage conditions to remove the affinity tag
- Biophysical studies to identify the purity and optimum buffer conditions to stabilize the expressed protein and estimation of the secondary structure
- Crystallization experiments, diffraction data collection, data processing, structure solution, structure refinement and deposition of three-dimensional structure in the protein data bank
- Production of optimum sized microcrystals for mix-and-diffuse tape drive serial crystallography experiments to investigate the catalysis of substrates
- Identification of the structural basis of thermostability and comparison of NhGH11 structure with homologous thermostable enzymes
- Identification of NhGH11 optimal parameters for the enzymatic reaction. Studies on biochemical parameters including the effect of temperature, pH, the effect of metal ions and chemical reagents

- Identification of NhGH11 specificity towards different substrates i.e. xylan beechwood, xylan polysaccharide, wheat arabinoxylan, azo-xyloglucan and mannan polysaccharide. Determination of Michaelis-Menten kinetics, Michaelis constant (K_m), turn over number (k_{cat}), and maximum rate of reaction (V_{max}) for different substrates
- Inhibition studies of NhGH11 in the presence of lignin derived phenolic compounds

2. Experimental part

2.1 General chemicals

2.1.1 Material and manufacturer

All chemicals were purchased in analytical grade from companies, Sigma Aldrich (Steinheim, Germany), Carl Roth (Karlsruhe, Germany) Merck (Darmstadt, Germany) and Megazyme (Megazyme International Ireland Limited, Ireland) and were used according to the manufacturer's instructions. All media and buffers were prepared in distilled water while the ultra-pure water (PureLab®flex, ELGA LabWater, Great Britain) was used in molecular biology assays.

Table 2.1 Consumables

Description	Manufacturer/Supplier
96-well vapor diffusion plates	Douglas Instruments Limited, UK
96-well SWISSCI MRC2 plates	Hampton Research Corp., USA
48-well SWISSCI MRC plates	Hampton Research Corp., USA
24-well linbro plates	Jena Bioscience GmbH, Germany
Amicon® Ultra 4, Ultra 15, Ultra 0.5, MWCO 3 kDa, 10 kDa	Merck Millipore (Merck KGaA), Germany
Vivaspin® centrifugal concentrators Vivaspin 15R, Vivaspin 500, 3 kDa MWCO	GE Healthcare Europe GmbH, Germany
Cellulose acetate membrane filter 0.2 µM	VWR International LLC, USA
Coverslips	VWR International LLC, USA
Ni-NTA resin	Macherey-Nagel GmbH & Co. KG, Germany
Amylose resin	New England Biolabs, 240 County Road, Ipswich, MA, USA

2. Experimental part

Syringes 1 mL, 2 mL, 5 mL, 10 mL, 20 mL, 50 mL	VWR International LLC, USA
Syringe filter 0.22 μm , 0.45 μm	VWR International LLC, USA
ZipTip®C18 Pipette Tips	Merck Millipore (Merck KGaA), Germany

Table. 2.2 Instruments

Device	Instrument	Manufacturer/Supplier
Agarose gel electrophoresis chamber	Blue Marine 100	Serva Electrophoresis GmbH, Germany
	Perfect Blue Gel system Mini	PEQLAB Biotechnology GmbH, Germany
Beamlines	P11	DESY, PETRA III, DESY, Hamburg, Germany
	P13	EMBL, PETRA III, DESY, Hamburg, Germany
CD spectrometer	J-815 CD	Jasco Inc., USA
Centrifuges	5415R/ 5415C/ 5804R/ 5810R Mini spin Plus	Eppendorf AG, Germany
	Multifuge X3R	Thermo Fisher Scientific Inc., USA
Crystallization robots	Honeybee 961	Genomic Solutions Inc., USA
	Oryx 4	Douglas Instruments Limited, UK
Crystal plate incubator 4 °C and 20 °C	RUMED 3001 incubator	Rubarth Apparate GmbH, Germany
DLS Device	SpectroLight 300	Xtal Concepts GmbH, Germany
	SpectroLight 600	Xtal Concepts GmbH, Germany

2. Experimental part

FPLC	ÄKTA Purifier P-901	GE Healthcare, USA
Freezer -20 °C	Liebherr premium	Liebherr International S.A., Switzerland
Freezer -80 °C	B35-85	FRYKA-Kältetechnik GmbH, Germany
Imaging	Leica TCS SP8 Confocal	Leica Microsystems GmbH, Germany
Incubator 37 °C	BD 56	BINDER GmbH, Germany
Incubation shaker	Innova 4330	New Brunswick Scientific Co., Inc. (Eppendorf AG), Germany
Magnetic stirrer	VMS-A	VWR International LLC, USA
Microbalance	Sartorius TE3102S CP2245- OCE	Sartorius AG, Germany
pH meter	Seven easy	Mettler-Toledo Inc., USA
Spectrophotometer	Nanodrop 2000c Nanodrop Lite	Thermo Fisher Scientific Inc., USA
SDS-PAGE	EV 231 (Power supply)	PEQLAB Biotechnology GmbH, Germany
Thermomixer	Comfort	Eppendorf AG, Germany
UV-light source	Crystal LIGHT 100	Nabitec GmbH, Germany
Vortex mixer	VF2	IKA Werke GmbH & Co. KG, Germany
X-ray source imaging plate X- ray detection System	I μ S	Mar research GmbH, Germany
SDS-PAGE	Four gel caster (SE275) Deluxe mini electrophoresis unit	Hoefer Inc., USA

2.2 Mutations and truncation of *N. haematococca* xylanase (NhGH11)

The sequence of NhGH11 (GenBank ID: GG698899.1, Uniprot ID: C7YSL3) was analyzed by applying the SignalP 4.1 server (<http://www.cbs.dtu.dk/services/SignalP-4.1/>) to identify and truncate the signal peptide sequence. The surface entropy reduction prediction server (<http://services.mbi.ucla.edu/SER/>) was applied to identify high entropy surface exposed lysines K4, K107, K108 and K168, after identification they were subsequently mutated to alanines to support crystallization experiments. A gene was synthesized by the company BioCat (Heidelberg, Germany). Codon optimization was done for the *E. coli* expression host. NdeI and HindIII were selected as 5' and 3' restriction sites (Fig. 2.1). A tobacco etch virus (TEV) protease cleavage site (ENLWFQG) was introduced between the maltose binding protein and a gene. Finally, a corresponding gene was ligated into the vector pMAL-c5X by Gibson assembly. Active site residues (E89 and E180) of NhGH11 were mutated to slow down enzymatic activity for soaking experiments. The construct with active site mutations was also synthesized by the company BioCat (Heidelberg, Germany), following the same cloning procedure.

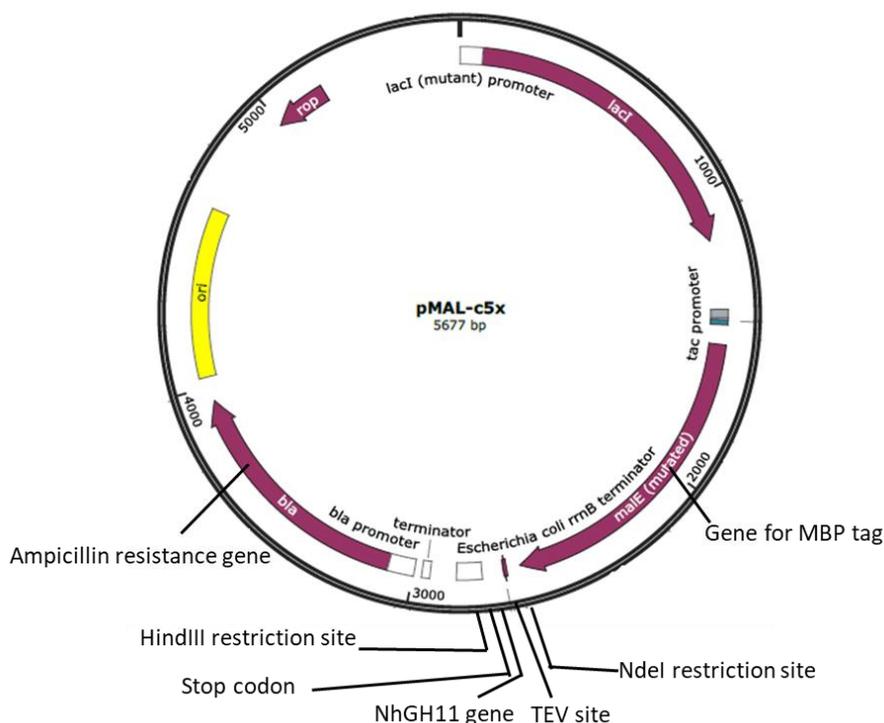


Fig. 2.1 Vector for *E. coli* expression of NhGH11. Vector contained ampicillin resistance gene (bla) and MBP gene (male). TEV cleavage site was introduced before NhGH11. NdeI and HindIII were selected as restriction sites. Figure source with modification: <https://www.addgene.org/67004/>.

2.3 Transformation of competent *E. coli* cells

Preparation of competent cells: chemically competent *E. coli* cells were prepared from different strains (Table 2.3). Cells of each strain were inoculated to 5 mL LB growth media and allowed to grow overnight at 37 °C and 180 rpm. After 12-14 hours, the culture was transferred to 50 mL growth media and allowed to grow until the optical density OD₆₀₀ of cells reached to 0.5. Afterwards, the cell culture was transferred to a 50 mL falcon tube, incubated on ice for 15 min and then centrifuged for 10 min at 4 °C. The supernatant was discarded and the cell pellet was washed three times with 5 mL of ice cold 100 mM CaCl₂ solution and finally resuspended in 1 mL of ice cold 100 mM CaCl₂ solution containing 10% glycerol. Aliquots of 50 µL cell suspensions were prepared on ice and flash frozen in liquid nitrogen to be stored at -80 °C.

Transformation: an aliquot of frozen competent *E. coli* cells was thawed on ice and 1 µL plasmid from 100 ng stock was added to the competent *E. coli* cells. Cells were gently mixed with the plasmid and incubated on ice for 15 min. Heat shock was applied to the cells at 42 °C for 45 s then they were immediately incubated on the ice for 5 min. Five hundred microliters of autoclaved SOC media (Table 2.4) was added and the cells were incubated at 37 °C and 350 rpm for one hour. The cells were centrifuged for 1 min at 2000 rpm and 350 µL of the supernatant was removed. The remaining suspension was gently resuspended and inoculated on agar plates containing 1% w/v agar in LB media and 100 µg/mL of ampicillin. Plates were incubated overnight at 37 °C.

Table 2.3 Bacterial Strains

Strain	Genotype	Supplier
DH5α	F ⁻ <i>endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20</i> φ80 <i>dlacZ</i> M15 (<i>lacZYA-argF</i>) U169, <i>hsdR17 (rK⁻ mK⁺)</i> , λ ⁻	Invitrogen (Thermo Fisher Scientific Inc.), USA
BL21 (DE3)	F- <i>ompT hsdSB (rB-mB-) gal dcm rne131</i> (DE3)	Invitrogen (Thermo Fisher Scientific Inc.), USA
BL21 (DE3) Star	F- <i>ompT hsdSB (rB-mB-) gal dcm rne131</i> (DE3) pLysS (CamR)	Invitrogen (Thermo Fisher Scientific Inc.), USA

2.4 Gene expression

One colony of transformed *E. coli* strain BL21 (DE3) was inoculated to 100 mL LB medium containing 100 µg/mL of ampicillin. The preculture was incubated overnight at 37 °C and 180 rpm and expression tests were performed, utilizing the TB, LB and autoinduction expression media (Table 2.4). Afterwards, the protein expression was initiated with 50 mL preculture in LB broth medium containing 100 µg/mL of ampicillin. The preculture was allowed to grow overnight at 37 °C, after approx. 14 hours, it was transferred to 1 L autoinduction medium. The cell culture was incubated at 37 °C until the optical density (OD₆₀₀) of cells reached 0.6, after that the temperature was reduced to 16 °C for overnight expression of NhGH11. The cell culture was transferred to a centrifugation bucket and centrifuged at 4000 × g and 4 °C for 30 min. The supernatant was discarded and the cell pellets were resuspended in the lysis buffer (Table 2.5). The resuspended cells were decanted into a falcon tube and stored at -20 °C.

Table 2.4 Medium and Buffers

Media	pH	Components	Concentration
SOC media	~7.2	Tryptone	2% (w/v)
		Yeast extract	0.5% (w/v)
		NaCl	10 mM
		KCl	2.5 mM
		MgCl ₂	10 mM
		MgSO ₄	10 mM
LB broth (Lennox)	~7.0	NaCl	5.0 g/L
		Tryptone	10 g/L
		Yeast extract	5.0 g/L
TB broth	~7.2	Tryptone	12 g/L
		Yeast extract	24 g/L
		K ₂ HPO ₄	12.5 g/L
		KH ₂ PO ₄	2.3 g/L
Autoinduction media	~7.2	Tryptone	10 g/L
		Yeast extract	5 g/L
		NaCl	5 g/L
		(NH ₄) ₂ SO ₄	3.3 g/L

2. Experimental part

	NH ₄ Cl	5.3 g/L
	KH ₂ PO ₄	6.8 g/L
	Na ₂ HPO ₄	7.1 g/L
	Glycerol	5 mL/L
	Glucose	0.5 g/L
	Lactose	2 g/L

2.5 Cell lysis and affinity chromatography

The cell lysis was done by sonication for 30 s and 30 s intervals in between; the procedure was repeated eight times. Afterwards, the cells were centrifuged at 4 °C and 17000 x g for 30 min to remove cell debris. Purification was done via affinity chromatography, utilizing the affinity of amylose resin with maltose binding protein (MBP). Purification was carried out by applying a standard affinity chromatography protocol¹⁵¹. Buffers applied for affinity purification are summarized in table 2.5. The steps of purification are shown in Fig. 2.2.

Step 1: according to the manufacturer's instructions, the amylose matrix was stored in 20% ethanol. The storage solution was removed and the column was washed with 8-10 column volumes of water followed by applying washing buffer.

Step 2: cell lysate was incubated with the amylose resin for 30 min at 8 °C.

Step 3: the unbound protein was removed and the column was washed with 8-10 column volumes of washing buffer.

Step 4: the bound protein was eluted with 10 mM maltose sugar included in the binding buffer. The absorbance at 280 nm was observed for each 2 mL fraction during the elution of NhGH11 from the column.

Step 5: the column was washed with 8-10 column volumes of washing buffer followed by applying water and then stored in 20% ethanol at 8 °C.

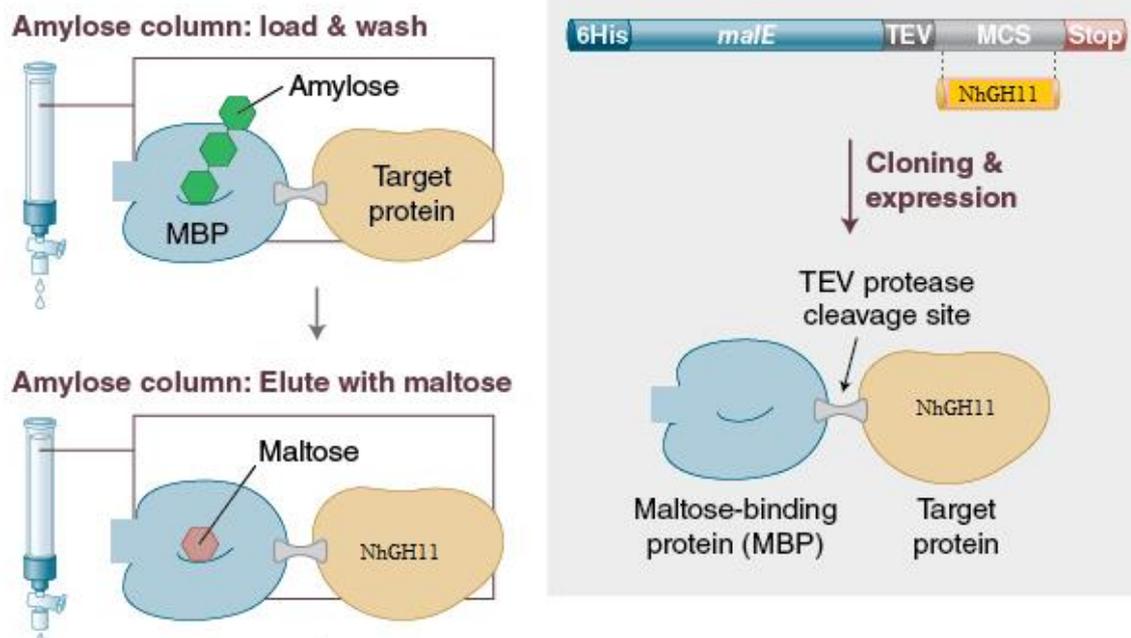


Fig. 2.2 Schematic diagram of affinity chromatography. Maltose binding protein (MBP) binds to amylose resin (green color). The column was washed to remove the unbound proteins. NhGH11 was eluted by maltose sugar. TEV protease cleavage site was introduced between MBP and NhGH11 to remove the MBP tag after affinity purification. Figure adapted from Website of New England BioLabs, source: <https://www.neb-online.de/proteinexpression/nebexpress-mbp-fusion-and-purification-system/> with modifications.

Table 2.5 Buffers used for purification

Name of buffer	Composition	Concentration
Lysis/Binding/ Washing Buffer	Tris-HCl, pH 7.4	20 mM
	NaCl	200 mM
	EDTA	0.5 mM
	DTT	0.5 mM
Elution Buffer	Tris-HCl, pH 7.4	20 mM
	NaCl	200 mM
	Maltose sugar	10 mM

2.6 Protein quantification

The protein concentration was estimated utilizing the absorption at 280 nm, the wavelength at which aromatic amino acids; mainly tryptophans have maximum absorption. In this context protein concentrations were measured by UV absorption at 280 nm using the molar absorption coefficient $58330 \text{ M}^{-1}\text{cm}^{-1}$, calculated by ProtParam server (<https://web.expasy.org/protparam>) for 20800 Da molecular weight of NhGH11. According to Beer's law, the concentration "c" [moles/L] is derived from the quotient of the absorption "A", which is a wavelength-dependent extinction coefficient [$\text{M}^{-1}\text{cm}^{-1}$] and the product of the extinction coefficient ϵ and the path length b [cm]:

$$c = \frac{A}{\epsilon * b} \quad (2.1)$$

Two microliters of protein samples were used for the measurements, applying the NanoDrop spectrophotometer (Thermo Fisher Scientific, USA).

2.7 Expression and purification of TEV protease

Preculture (50 mL LB medium) was inoculated by glycerol stock of transformed *E. coli* strain BL21 (DE3) star pre-transformed with the vector containing TEV gene. The antibiotics, 100 μL chloramphenicol (stock conc. 30 mg/ml) and 100 μL ampicillin (stock conc. 100 mg/ml) were included. Preculture was incubated overnight at 37 °C and then transferred to 1 L LB medium containing antibiotics mentioned above. The optical density (OD_{600}) was measured after regular intervals and protein expression was induced with 1 mM final concentration of Isopropyl β - d-1-thiogalactopyranoside (IPTG) solution at the cell density of 0.5. Protein expression was done at 37 °C for 4 hours. Afterwards, the culture was transferred to centrifugation buckets and centrifuged for 30 min at 4 °C and 4000 x g. The supernatant was discarded and the cell pellets were resuspended in the lysis buffer. The cell lysis was done by sonication for 30 s and 30 s intervals in between; the procedure was repeated eight times. Afterwards, the cells were centrifuged at 4 °C and 17000 x g for 30 min to remove cell debris. The clear lysate was transferred to the Ni-NTA column. Buffers used for the purification of TEV are summarized in table 2.6. The matrix was washed with seven column volume of washing buffer followed by elution. Elution fractions were collected from each elution buffer E1-E5. The column was washed with five column volume of E6 buffer. Elution fractions were pooled into the dialyzing membrane and dialyzed overnight against the dialysis buffer. A TEV protease was concentrated to 1 mg/mL.

The concentration was measured by UV absorption at 280 nm using the molar absorption coefficient $31970 \text{ M}^{-1}\text{cm}^{-1}$, for 28533 Da molecular weight of TEV protease. The TEV protease was stored at $-20 \text{ }^{\circ}\text{C}$ for further use. The affinity-purified NhGH11 was incubated with 1 mg of TEV protease/10 mg of NhGH11 incubated overnight at $8 \text{ }^{\circ}\text{C}$. After cleavage, the TEV protease was removed from the MBP+NhGH11 mixture by using a Ni-NTA matrix and the flow-through obtained from the column was collected for further purification, applying size exclusion chromatography.

Table 2.6 Buffers used for purification of TEV

Name of buffer	Composition	Concentration
Binding/Washing Buffer	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	25 mM
Elution Buffers		
E1	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	50 mM
E2	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	100 mM
E3	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	150 mM
E4	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	200 mM

E5	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	250 mM
E6	Na ₂ HPO ₄ , pH 8.0	20 mM
	NaCl	200 mM
	Glycerol	10%
	Imidazole	1M
Dialysis Buffer	Na ₂ HPO ₄ , pH 8.0	20 mM
	EDTA	0.5 mM
	Glycerol	30%

2.8 Sample concentration

Protein solutions were concentrated up to 5 mL with approx. concentration of 10 mg/mL by applying Vivaspin centrifugal concentrators with 3 kDa molecular weight cutoff (GE Healthcare Europe GmbH, Germany).

2.9 Size exclusion chromatography (SEC)

The ÄKTA purifier was used for size exclusion chromatography. The Hiload 16/60 Superdex 75 SEC column (GE Life Sciences) was washed with two column volumes of dH₂O, followed by equilibration with one column volume of SEC buffer (20 mM Tris, 200 mM NaCl at pH 7.4). The protein sample was concentrated to a final volume of 5 mL to be injected into the column. The flow rate was set at 0.7 mL/min and the absorption was monitored at 280 nm. Elution fractions were analyzed by SDS-PAGE.

2.10 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Buffers and solutions used for SDS-PAGE are summarized in table 2.7. The main component of SDS gel is acrylamide, which is polymerized by the addition of tetramethyl ethylenediamine (TEMED) and ammonium persulfate (APS). The gel consisted of two different layers; a narrow pore separating gel and a coarser stacking gel. Stacking gel captured the proteins between the glycine and chloride anions. The stacking gel consisted of acidic pH of 6.8 where the chloride anions flow faster than the glycine anions, causing the proteins to be focused between the anions. When the proteins enter the separating gel consisted at pH 8.8, the proteins are no longer focused between the

two anions and separated by molecular weight size¹⁵².

Protein samples of NhGH11 were mixed with 2x SDS-sample buffer before denaturing them at 96 °C for 5 min. The samples were spin down briefly and 10 µL samples were loaded to the gel. The 1x electrode buffer was used to run gels in a vertical gel electrophoresis chamber, applying 120 W power and 25 mA current. After approx. one hour, the gels were stained in a coomassie staining solution with continuous shaking for 30 min. Afterwards, the gels were destained until the protein bands can be clearly distinguished from the background.

Table 2.7 Buffers and solutions used for SDS-PAGE

Gel type	Component	Volume
Stacking Gel (4%)	ddH ₂ O	9.2 mL
	30% acrylamide/bisacrylamide	2.0 mL
	(37.5:1)	3.8 mL
	Stacking gel buffer	150 µL
	10% (w/v) SDS	15 µL
	TEMED	75 µL
Separating Gel (12%)	ddH ₂ O	10.2 mL
	30% acrylamide/bisacrylamide	12.0 mL
	(37.5:1)	7.5 mL
	Stacking gel buffer	300 µL
	10% (w/v) SDS	15 µL
	TEMED	150 µL
Coomassie Staining Solution	Isopropanol	25% (v/v)
	Acetic Acid	10% (v/v)
	Coomassie Brilliant Blue	0.1% (w/v)
Coomassie Destaining Solution	Acetic Acid	20% (v/v)

2 x sample buffer SDS PAGE	Tris-HCl, pH 6.8	20 mM
	SDS	4% (w/v)
	Glycerol	20% (v/v)
	DTT	50 mM
	Bromophenol blue	0.04% (w/v)
Separating Gel Buffer	Tris-HCl, pH 8.8	1.5M
	SDS	4% (w/v)
Stacking Gel Buffer	Tris-HCl, pH 6.8	5 M
	SDS	4% (w/v)
10x Electrode Buffer	Glycine	1.9M
	Tris-HCl	250 mM
	SDS	1% (w/v)

2.11 Dynamic light scattering

Dynamic light scattering (DLS) was applied to analyze the hydrodynamic radius (R_h) of NhGH11. Monochromatic light is exposed to the protein molecules in solution and the light is scattered in all directions. The positions of the scatter origins varied all the time because of the Brownian motion of particles and fluctuations of scattering intensities. These fluctuations provided information about the velocity of the particles and the diffusion coefficient. The hydrodynamic radius (R_H) is derived according to the Stoke's-Einstein-equation¹⁵³:

$$R_H = \frac{kT}{6\pi\eta D} \quad (2.2)$$

Where k is the Boltzmann's constant [J/K], T is the temperature [K], η is the viscosity of the solution [Ns/m²] and D is the diffusion coefficient [m²/s].

DLS measurements were performed, applying the SpectroSize 300 (Xtal Concepts GmbH, Germany) using a 70 μ L protein solution in a cuvette. Before measurements, protein samples were centrifuged at 12000 x g for at least 30 min.

2.12 Circular dichroism (CD) spectroscopy

CD spectra provide information about the secondary structure and folding states of proteins. The optical activities of proteins are different so circularly polarized light is absorbed differently depending on the composition. The α -helices and β -sheets showed

certain absorption minima and maxima at distinct wavelengths and hence the relative content of these secondary structure elements was estimated¹⁵⁴.

Circular dichroism (CD) spectroscopy was used to perform thermostability analysis by thermal denaturation of NhGH11 using a JASCO J-815 CD spectrometer (Jasco, UK). For CD investigations, a solution of NhGH11 was diluted to 0.12 mg/mL in 5 mM NaF solution. The spectra were recorded from 20-90 °C with an increment of 1 °C/min. The measurements were done in a quartz cuvette (1 mm path length) and the wavelengths applied were ranging from 190 up to 260 nm. The ellipticity was plotted against the wavelengths and temperatures, using the software J-815 Spectra manager (Jasco, UK). Ten CD spectra were recorded after 10 °C intervals. The corresponding buffer spectra were recorded for background subtraction. The identification of secondary structure elements was done following standard curves of reference proteins¹⁵⁵. Ellipticity can be reported by using the equation given below:

$$\theta = \frac{180 \cdot \ln 10}{4\pi} (ER - EL) \quad (2.3)$$

Ellipticity is defined as the angle whose tangent is the ratio of the minor to the major axis of the ellipse. Change in ellipticity is $\Delta E = (ER - EL)$ while ER and EL are the absorbance of right and left-handed circular polarized light¹⁵⁶.

2.13 Crystallization

Initial crystallization screening was performed by vapor diffusion-hanging drop using NhGH11 (10 mg/mL) and different concentrations of ammonium sulfate. Initial inter-grown crystals with approx. dimensions between 50-100 μm were obtained at 20 °C after two days, by mixing equal volumes of precipitant solution (1 M ammonium sulfate, 100 mM sodium citrate pH 5.5) and protein solution. To obtain uniform size crystals streak seeding technique was applied.

Seed stock: To prepare seed stocks initial crystals obtained were selected from hanging drops and crushed under a stereomicroscope, using a crystal crusher tool (Hampton research). The reservoir solution (10 μL) was pipetted to the drop and the crystal crusher was washed in the drop. Remnants of seed stock were collected by washing the drop with reservoir solution (10 μL). Seed stock was transferred to the tube containing seed bead (Molecular Dimensions Ltd., USA), vortexed three times for 30 s with an interval of 30 s on ice and stored at -80 °C.

Streak seeding technique: hanging drop experiments were set up by mixing equal volumes of precipitant solution (1 M ammonium sulfate, 100 mM sodium citrate, pH 5.5) and 10 mg/mL of NhGH11. The seed stock was thawed on ice and streak seeding was applied with an artificial hair briefly dipped in the seed stock and the tip of the hair was moved in a straight line across the drop containing the protein and precipitant. Uniform-sized crystals with dimensions between 100-150 μm were obtained after two days.

2.13.1 Diffraction data collection

Crystals obtained by streak seeding conditions were applied for diffraction data collection. Firstly, cryoprotectant solutions were tested applying a range of glycerol concentrations from 15-30%. Precipitant solution with 25% (v/v) glycerol was found suitable for data collection. Therefore, before the data collection, the crystal was treated with a cryoprotectant solution consisting of the reservoir solution supplemented with 25% glycerol. Afterwards, a crystal was mounted in a cryo nylon loop (Mounted CryoLoop, Hampton Research, US). The first dataset was collected using the in-house x-ray source mar345 fast image plate detector (MARRESEARCH, Germany). Few images were collected and indexed to determine the phi range for completeness. The parameters for data collection consisted of an oscillation range of 1° (180 images in total), exposure time of 60 s and 200 mm detector to crystal distance. A full dataset was collected to 2.00 \AA resolution.

The beamline P13 (EMBL, Hamburg) at PETRA III, DESY, was used to collect the high resolution data. A crystal was flash frozen in liquid nitrogen and measurements were conducted under cryogenic conditions at -173°C using a liquid nitrogen stream after applying cryoprotectant solution. The data collection strategy included a detector to crystal distance of 138.67 mm, an exposure time of 0.0377440 s/frame, a wavelength of 0.97 \AA and an oscillation range of 0.1° . In total 1800 images were collected and a complete dataset was collected to atomic resolution (1.00 \AA). The dataset was processed applying the XDS package¹⁵⁷. The phase problem was solved by applying molecular replacement using the homologous structure of *F. oxysporum* (PDB code: 5JRM), having 82% sequence homology to NhGH11.

2.13.2 Tape-drive experiments

Preparation of microcrystals: crystals obtained by the vapor diffusion-hanging drop method with precipitant containing 1 M ammonium sulfate, 100 mM sodium citrate pH 5.5, was selected to prepare a seed stock (seed stock preparation is described in section 2.13). Protein solution 15 mg/mL, precipitant solution and seed stock were mixed with ratio 1:1:0.5. The mixture was vortexed for 30 s applying ten minutes intervals. After 30 min the microcrystals were centrifuged at 200 rpm and the supernatant was replaced with a precipitant solution. Applying the same protocol, microcrystals were obtained under two precipitant conditions.

Precipitant 1: 1M (NH₄)₂SO₄, 100 mM sodium citrate pH 5.5.

Precipitant 2: 200 mM (NH₄)₂SO₄, 100 mM sodium citrate pH 5.5 and 20% PEG 6000. Microcrystals obtained under both precipitant conditions were tested for diffraction data collection.

Experimental setup: The tape-drive experiments were set up as described before¹²⁰. The experiments were performed at beamline P11, PETRA III DESY, by Dominik Oberthür and his team. A brief description of the experimental setup includes, a rotating beam chopper was placed in the path of the X-ray beam to generate repeated pulses of 4.73 ms duration and 25 Hz rate. A Pilatus 6M detector was applied to collect each single diffraction pattern. A fused silica capillary with internal two openings for crystal suspension and substrate solution was used to deposit a thin stream of crystals on a polyimide tape. The tape speed was set to 1 mm/s allowing a 70-80% hit rate during the experiments. Data collection was done using mixing times of 1, 2, 3, 4, 5 and 7 s, applying microcrystals and a substrate (70 mM xylopentose prepared in precipitant solution) suspension with a mixing ratio of 1:1. The collected data were indexed and integrated by Dominik Oberthür using the program CrystFEL version 0.6.1. The integrated reflection file with mtz format was further applied for refinement by applying the program PHENIX 1.8.4_1496.

2.13.3 Soaking experiments

The crystals obtained from two precipitant conditions mentioned in table 2.8 were separately soaked with substrates: xylobiose (X2), xylotriose (X3) and inhibitors: cinnamic acid, coumaric acid, ferulic acid and caffeic acid. The summary of soaking experiments is shown in Fig. 2.3, and table 2.8. Crystals from two precipitant conditions

were used for soaking; precipitant 1: 1M $(\text{NH}_4)_2\text{SO}_4$, 100 mM sodium citrate pH 5.5 and precipitant 2: 200 mM $(\text{NH}_4)_2\text{SO}_4$, 100 mM sodium citrate pH 5.5, 20% PEG 6000.

Soaking of crystals in substrates: stock solutions (150 mM) of substrates were prepared in precipitant 1 and 2 separately, but the low molecular weight PEG 400 was used instead of PEG 6000 for precipitant 2 to reduce the viscosity of solutions.

Three types of crystals were used for soaking of crystals in substrates:

Crystals of type 1: NhGH11 crystals grown in precipitant condition 1.

Crystals of type 2: NhGH11 crystals grown in precipitant condition 2.

Crystals of type 3: NhGH11 active site mutant crystals grown in precipitant condition 1.

NhGH11 active site mutant crystals grown in precipitant condition 1 were used for soaking with substrates, however, NhGH11 active site mutant crystals could not grow in precipitant condition 2. Crystallization, data collection, and structure of NhGH11 active site mutant structure are shown in the supplementary information, S. 4.4 and table 4.1 respectively.

All three types of crystals were transferred to new drops, each containing the substrate and the cryoprotectant (25% glycerol) prepared separately for each precipitant (Fig. 2.3, Method 1). Soaking was performed in a gradient of substrate concentration starting from 10 mM to 75 mM, all soaked crystals were used for data collection. In table 2.6 the highest soaking concentration is mentioned along with the soaking time. Crystals were flash-frozen in liquid nitrogen after soaking for 30s, 40s, 50s, 60s, 2 min, 3 min, 4 min and 5 min.

Soaking of crystals in Inhibitors: stock solutions (100 mM) of inhibitors were prepared in 50% (v/v) methanol and preheated sodium acetate buffer, pH 6.0. Further dilutions were done in the precipitant 1 and 2. Two types of crystals used for soaking substrates:

Crystals of type 1: NhGH11 crystals grown in precipitant 1

Crystals of type 2: NhGH11 crystals grown in precipitant 2

Crystals were transferred to new drops, each containing one inhibitor. Drops were closed for overnight soaking (Fig. 2.3, method 2). Cryoprotectant was prepared in the same precipitant condition 1 or 2 containing 25% glycerol. The concentrations of inhibitors applied for soaking experiments were also applied in cryoprotectant solutions that were applied before flash freezing in liquid nitrogen.

2. Experimental part

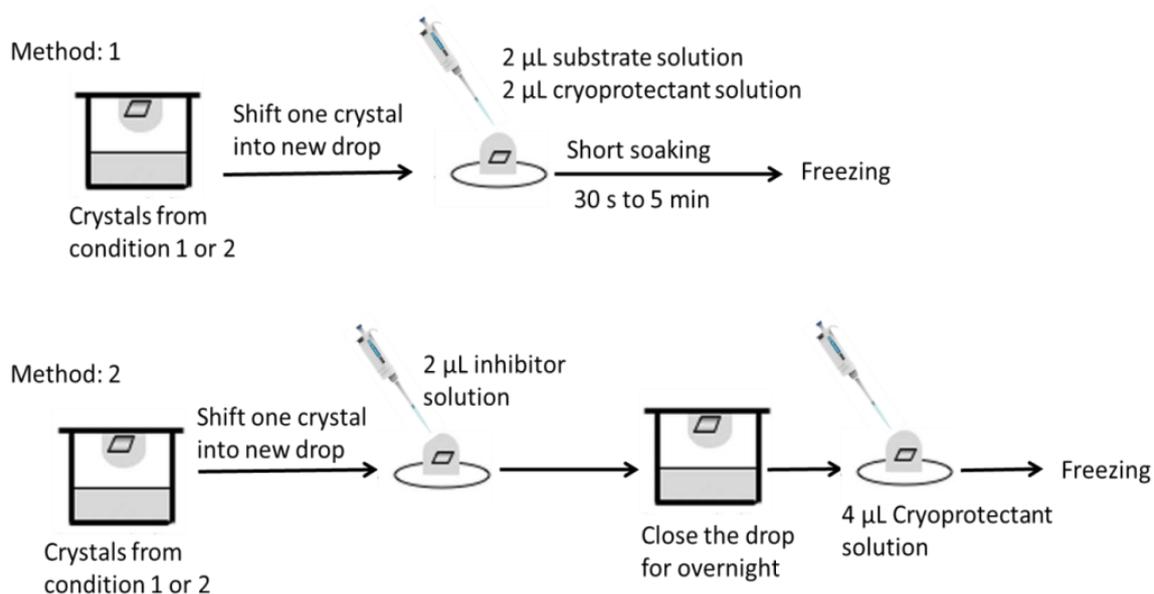


Fig. 2.3 Schematic diagram of soaking. Method 1: soaking of NhGH11 crystals with substrates. Method 2: soaking of NhGH11 crystals with inhibitors.

Table 2.8 Summary of soaking experiments

Crystals used for soaking	Soaking Time	Substrates/ Inhibitor	Concentration
Crystals of type 1 and 3	30 s	150 mM stock solution of X3 dissolved in 1M (NH ₄) ₂ SO ₄ , 100 mM sodium citrate pH 5.5	75 mM
	40 s		
	60 s		
	2 min		
	3 min		
	4 min		
	5 min		
Crystals of type 2	30 s	150 mM stock solution of X3 dissolved in PEG 400 (20%), 100 mM sodium citrate pH 5.5	75 mM
	40 s		
	60 s		
	2 min		
	3 min		
	4 min		
	5 min		
Crystals of type 1 and 3	30 s	150 mM stock solution of X2 dissolved in 1M (NH ₄) ₂ SO ₄ , 100 mM sodium citrate pH 5.5	75 mM
	40 s		
	60 s		
	2 min		
	3 min		
	4 min		
	5 min		

2. Experimental part

Crystals of type 2	30 s	150 mM stock solution of X2 dissolved in PEG 400 (20%), 100 mM sodium citrate pH 5.5	75 mM
	40 s		
	60 s		
	2 min		
	3 min		
	4 min		
	5 min		
Crystals of type 1 and 2	Overnight	Coumaric acid	2 mM
		Ferulic acid	2 mM
		Cinnamic acid	2 mM
		Caffeic acid	2 mM
Crystals of type 1 and 2	Overnight	Coumaric acid	4 mM
		Ferulic acid	4 mM
		Cinnamic acid	4 mM
		Caffeic acid	4 mM
Crystals of type 1 and 2	Overnight	Coumaric acid	7 mM
		Ferulic acid	7 mM
		Cinnamic acid	7 mM
		Caffeic acid	7 mM

2.14 Molecular docking

Superimposition of NhGH11 with homologous structure was done using the PDB model 4HK8 xylanase to predict the active site residues. The Flexidock program from the Sybyl-X program package (Tripos, USA) was used to dock xylohexose to the predicted active site of NhGH11. For the preparation of docking files, the refined model of NhGH11 was used, while the ligand complex coordinate file was prepared using the PDB model 4HK8. Hydrogen atoms were added to the ligand and charges were calculated both for the protein and the ligand. The ligand was properly positioned to the active site cavity of NhGH11. The free energy of docked model was minimized by the software. The active site of the docked complex was compared with the xylohexose bound PDB model 4HK8.

2.15 Biochemical characterization

Buffers and substrate solutions used in biochemical assays are summarized in table 2.9.

2.15.1 Effect of temperature and pH on enzyme activity

For optimum temperature: the reaction mixtures (NhGH11+substrate) were incubated at

different temperatures (30-80 °C) in 50 mM sodium citrate buffer at pH 5.5.

For optimum pH: the reaction mixtures (NhGH11+substrate) were incubated with a range of 100 mM MacIlvain buffers (pH 2-12).

For temperature-dependent stability: NhGH11 was incubated in 50 mM sodium citrate pH 6.0 at temperatures 40 °C, 45 °C and 50 °C for 8 hours. The samples were collected after each hour for activity assay.

Biochemical reactions for optimum temperature, optimum pH and temperature-dependent stability were completed by following the procedure given below:

DNS assay: Ten microliters of purified NhGH11 (0.16 mg/mL) was incubated with 40 µL of xylan beechwood (1%) and 50 µL of the buffer, for 15 min at the specified reaction temperature. The DNS solution 100 µL was mixed and the reaction mixture was boiled at 90 °C for 10 min and then cooled on ice. The absorbance of each sample was measured at 540 nm¹⁵⁸.

The standard xylose curve was prepared using the gradient of xylose sugar following the DNS assay. All reactions mentioned below were performed utilizing sodium citrate buffer pH 6.0 and reaction temperature 45 °C.

2.15.2 Effect of metal ions and chemical reagents

Stock solutions of 100 mM concentration of the salts (NaCl, KCl, CoCl₂, CaCl₂, MgCl₂, ZnSO₄, MnCl₂, FeSO₄ and CuCl₂) were prepared and further diluted with 50 mM sodium acetate buffer pH 6.0 to obtain 10 mM final concentrations. Stock solutions of chemical reagents (tween-80, triton-100, β-mercaptoethanol, EDTA, SDS) were prepared to a final concentration of 2% (w/v) except for sodium azide which was 0.2%. NhGH11 was incubated with salts and chemical reagents and the reactions were completed by following the procedure of DNS assay¹⁵⁸ described in section 2.15.1. The residual activity was calculated and data were represented as the mean of triplicates with the standard deviation.

2.15.3 Substrate kinetics

The catalytic activity of NhGH11 was tested towards xylan beechwood (Sigma, USA), xylan polysaccharide from *Palmaria palmata* (Elicityl, France), wheat arabinoxylan (Megazyme, Australia), azo-xyloglucan from tamarind (Megazyme, Australia) and mannan polysaccharide (Sigma, USA). The hydrolysis reactions with the fixed enzyme concentration (0.16 mg/mL) and substrates range from 0.05 up to 8.0 mg/mL, the reactions

were completed by following the procedure of DNS assay¹⁵⁸ described in section 2.15.1. The Michaelis Menten constant (K_m) and maximum velocity were calculated by using the software Origin Pro 2018, USA.

2.15.4 Identification of products

MALDI-TOF-MS was used to measure the molecular masses of fragmented ions originating from the hydrolysis of xylopentose (X5). NhGH11 was incubated with X5 at optimum temperature and pH conditions previously determined for NhGH11, however in a low ionic strength sodium citrate buffer (5 mM). After enzymatic hydrolysis, the solution was centrifuged at $16000 \times g$ for 30 min to remove any precipitated material. MALDI-TOF-MS measurements were performed using an Ultrafle Xtreme mass spectrometer (Bruker Daltonik, Bremen, Germany) equipped with a 2 kHz smart beam-II Laser and by using Dihydroxyacetophenone matrix. Data were automatically acquired using Flex control 3.0 and Maldi Biotyper Automation Control 2.0. Software (Bruker Daltonics GmbH, Bremen, Germany). The method of identification included the m/z from 100 to 800 Da. For each spectrum maximum of 100 peaks were considered and compared with the expected molecular weights of xylooligosaccharide with sodium ions.

2.16 Inhibition assays

2.16.1 Residual activity

Stock solutions (100 mM) of inhibitors: cinnamic acid, coumaric acid, ferulic acid and caffeic acid were prepared in 50% (v/v) methanol and preheated sodium acetate buffer, pH 6.0. Further dilutions to final concentrations (300 μ M to 4.5 mM) were also done in the same buffer. NhGH11 was incubated with inhibitors and the reactions were completed by following the procedure of DNS assay¹⁵⁸ described in section 2.15.1. A control test was performed in the absence of inhibitors. The residual activity was calculated and data were represented as the mean of triplicates with the standard deviation.

Table 2.9 Buffers and substrate solutions for biochemical assays

Buffers/solutions	Composition	Concentration
Sodium citrate buffer	$\text{Na}_3\text{C}_6\text{H}_5\text{O}_7 \cdot 2\text{H}_2\text{O}$, pH 6.0	50 mM
Sodium acetate buffer	$\text{C}_2\text{H}_3\text{NaO}_2$, pH 6.0	
MacIlvain Buffers	$\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ and $\text{C}_6\text{H}_8\text{O}_7$, pH 2-12	100 mM
Xylose	$\text{C}_5\text{H}_{10}\text{O}_5$	10 mM
DNS reagent	1% (w/v) DNS, 0.2% (v/v) phenol, 1% (w/v) NaOH, 0.002% glucose and 0.05% (w/v) Na_2SO_3	1%
Xylan beechwood	β -1,4 linked xylan with 8.7% of 4- <i>O</i> -methyl glucuronic acid and 5.7% other sugars, Sigma, USA	1%
Xylan polysaccharide	(β -1,3- and β -1,4-linked D-xylose polymer, Elicityl, France	1%
Wheat arabinoxylan	β -1,4-linked xylan with overall 36% of arabinose, 51% xylose, 6.5% glucose and 4.4% of mannose sugars, Megazyme, Australia	1%
Azo-xyloglucan	β -1,4 linked glucose with branches of galactose, xylose, arabinose and other sugars, Megazyme, Australia	1%
Mannan polysaccharide	β -1,4-linked mannose with 2% galactose and other sugars, Sigma, USA	1%
Xylopentose	β -1,4-linked 5 xylose sugar subunits	3 mg/mL
Coumaric acid	$\text{C}_9\text{H}_8\text{O}_3$ in 50% methanol	100 mM
Caffeic acid	$\text{C}_9\text{H}_8\text{O}_4$ in 50% methanol	100 mM
Cinnamic acid	$\text{C}_9\text{H}_8\text{O}_2$ in 50% methanol	100 mM
Ferulic acid	$\text{C}_{10}\text{H}_{10}\text{O}_4$ in 50% methanol	100 mM

2.16.2 Thermofluor assays

Phenolic compounds 50 mM stock solutions were prepared in 50% methanol, further dilutions to concentrations (150 μ M to 1.20 mM) were done with 10 mM preheated sodium acetate buffer pH of 6.0. NhGH11 (0.16 mg/mL) was incubated with final concentrations of inhibitors in a 96 well PCR plate. A freshly prepared stock solution (62x) of an SYPRO Orange fluorescent dye was added to each well having a final volume of 21 μ L. For negative control, the dye was mixed with the buffer containing inhibitors. The PCR plate was sealed with optical clear stick tape (Sigma Aldrich, Steinheim, Germany), the plate was centrifuged (4°C, 30 s, 2500 x g) and analyzed on a real-time PCR machine using a temperature gradient of 1 °C/min from 20 to 95 °C. The melting curves obtained from each well were analyzed by applying the software Graph prism 5.00.288, (California, USA). The temperature was plotted against the fluorescence to obtain a sigmoidal curve. Fluorescence was calculated by using the equation given below¹⁵⁹:

$$Y = Bottom + \frac{Top - Bottom}{1 + e^{(T_m - \frac{x}{slope})}} \quad (2.4)$$

Y: fluorescence emission (in arbitrary units), X: temperature in °C, Bottom: baseline of fluorescence at low temperatures (in arbitrary units), Top: maximal fluorescence, Slope: steepness of the curve, T_m: melting temperature of NhGH11 with phenolic inhibitors in °C.

2.16.3 Tryptophan fluorescence spectrophotometry

Phenolic compounds 50 mM stock solutions were prepared in 50% methanol, further dilutions to concentrations (0.3 to 1000 μ M) were done with 10 mM preheated sodium acetate buffer pH 6.0. Control measurements were performed for NhGH11 (0.16 mg/mL) without the addition of phenolic compounds in the same standard conditions.

Sodium acetate buffer 10 mM, pH 6.0 was used as a reference for control measurements. The total volume of 150 μ L of protein and inhibitor mixture was prepared in a 96 well plate. The excitation was done at 280 nm and the emission spectra were measured between 310 to 420 nm. Blank solutions were prepared in the buffer containing the phenolic compound. All measurements were performed in triplicates. The software Graph prism 5.00.288, (California, USA) was applied to analyze non-linear regression one-sided ligand binding equation¹⁶⁰ given below:

$$Y = \frac{B_{max} \times X}{K_d + X} \quad (2.5)$$

Y: relative fluorescence plotted on the axis, X: ligand concentration plotted on the axis,
B_{max}: the maximum achievable value of Y extrapolated to very high concentrations of X,
K_d: dissociation constant

3. Results and discussion

3.1 Surface entropy reduction prediction (SERp)

The amino acid sequence of *Nectria haematococca* xylanase from glycosyl hydrolase family 11 (NhGH11) was analyzed by the SERp server. The conformational entropy was computed using default parameters of SERp server to predict the surface exposed high entropy residues present in the secondary structure of NhGH11. The predicted surface exposed high entropy residues were divided into three clusters (cluster 1: K4, cluster 2: K107, K108, and cluster 3: K168) as shown in Fig. 3.1. These high entropy surface exposed lysines of NhGH11 were subsequently mutated to alanine residues to support crystallization experiments.

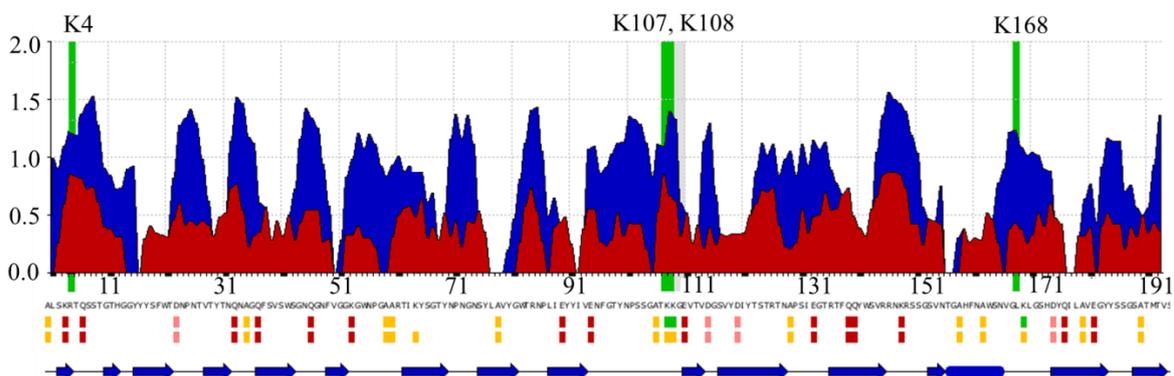


Fig. 3.1 Summary of high entropy scores by SERp server: red peaks correspond to surface entropy (arbitrary units), blue peaks indicate residues predicted to be in loops. Below the peaks, the high entropy residues not recommended for mutations are shown in pink while the proposed mutations (K4, K107, K108, and K168) are highlighted in green, all mutable residues are shown in red, and low entropy target residues are shown in yellow. Also, the predicted secondary structure elements are shown at the bottom in blue color.

In nature, proteins have evolved with a surface-entropy shield created by high entropy residues to hinder the phenomena of crystal formation inside the cells. Therefore, the high entropy residues on the surface of proteins also hinder the *in vitro* crystallization process due to the high entropy barrier^{161,162}. Consequently, the SERp server was developed to provide a higher propensity of crystal formation via mutation of high entropy residues i.e., surface exposed lysines, glutamates and glutamines¹⁶³. Predictions of SERp server enhanced the crystallization of several proteins by overcoming the high entropy barrier that needs to reach the labile zone of the phase diagram. Approximately, 40 mutants were tested for mutation of high entropy residues and alanine residues showed particularly the best replacement with the lowest entropy values^{164,165}.

3.2 Protein expression and purification

The recombinant protein expression of NhGH11 was optimized by using the *E. coli* strain BL21 (DE3) and autoinduction medium. Maltose binding protein (MBP) was used as an affinity tag for purification that improved the soluble protein fraction and the native folding of NhGH11. The tobacco etch virus (TEV) protease cleavage site was introduced between the MBP tag and NhGH11 as shown in Fig. 3.2. The affinity-purified NhGH11 was incubated with 1 mg of TEV protease/10 mg of fusion protein at 8 °C overnight. The TEV protease acted on a specific recognition sequence (ENLWFQG) and successfully cleaved the MBP tag from NhGH11. Afterwards, the TEV protease was removed from the MPB+NhGH11 mixture by Ni-NTA affinity resin and further purification of NhGH11 was done by size exclusion chromatography.

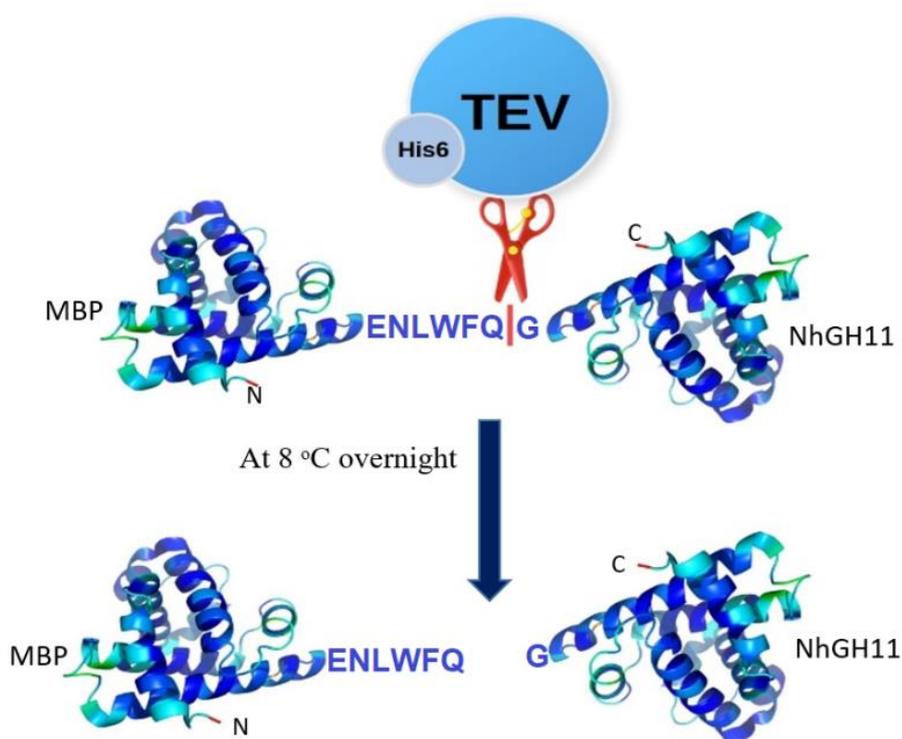


Fig. 3.2 Cleavage of MBP-NhGH11 with TEV protease. The TEV protease was expressed and purified with His-tag and was further used to hydrolyze the bond between Q|G at TEV cleavage site. Therefore, the MBP tag is separated from NhGH11. Source: <https://www.protean.bio>, with modifications.

Size-exclusion chromatogram, SDS gel and dynamic light scattering (DLS) results are shown in Fig. 3.3. The highly pure protein sample was monodispersed with radius $R_H = 2.3 \pm 0.03$ nm, which corresponds to approx. 22 kDa molecular weight, observed by DLS. The highly pure and monodispersed protein sample was further used for crystallization experiments. The above mentioned approach was adapted for NhGH11 to

overcome the expression and solubility problems observed for glycosyl hydrolases from family GH10 and AA9, briefly described by the flow sheet diagram in the supplementary information, S. 4.1.

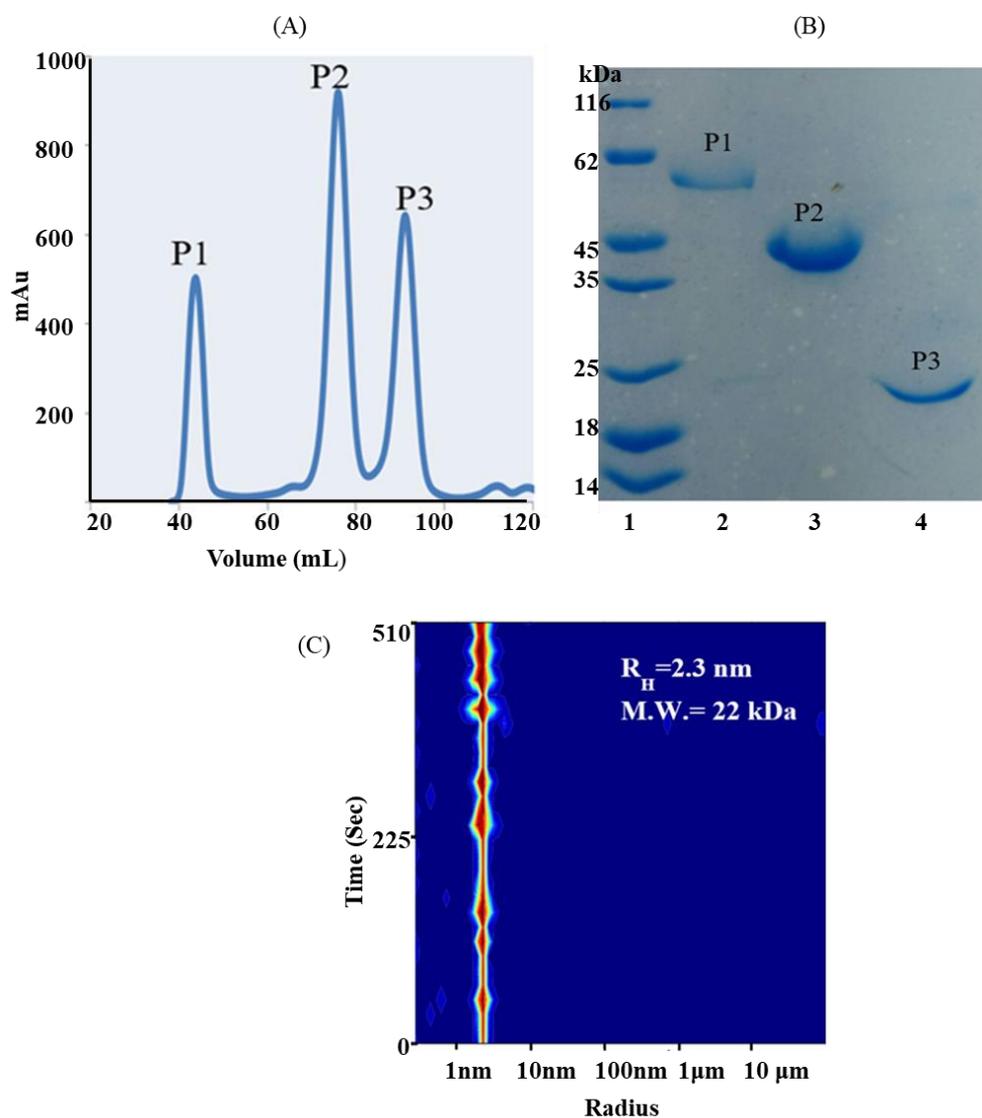


Fig. 3.3 Purification and physical characterization of NhGH11. (A) Size-exclusion chromatogram indicating peaks, P1: MBP-NhGH11, P2: MBP after cleavage, P3: NhGH11 pure after size exclusion chromatography. (B) SDS PAGE indicating lane 1: molecular marker, lane 2: MBP-NhGH11, approx. 60kDa, lane 3: MBP after cleavage, approx. 40 kDa, lane 4: NhGH11 after cleavage from MBP, approx. 22kDa. (C) Dynamic Light Scattering (DLS) of NhGH11 (10 mg/ml), showing a monodisperse protein solution, radius $R_H = 2.3 \pm 0.03$ nm, which corresponds to approx. 22 kDa molecular weight.

3.3 Crystallization experiments

Crystallization conditions of NhGH11 were successfully optimized by applying the vapor diffusion hanging drop method. Equal volumes of NhGH11 (10 mg/mL) and precipitant solution (1 M ammonium sulfate, 100 mM sodium citrate, pH 5.5) were mixed. The streak seeding technique was applied to introduce micro seeds into a hanging drop. Uniform-sized crystals were obtained with approx. length of 100 μm as shown in Fig. 3.4 A and a phase diagram of crystallization for the vapor diffusion method is shown in Fig. 3.4 B.

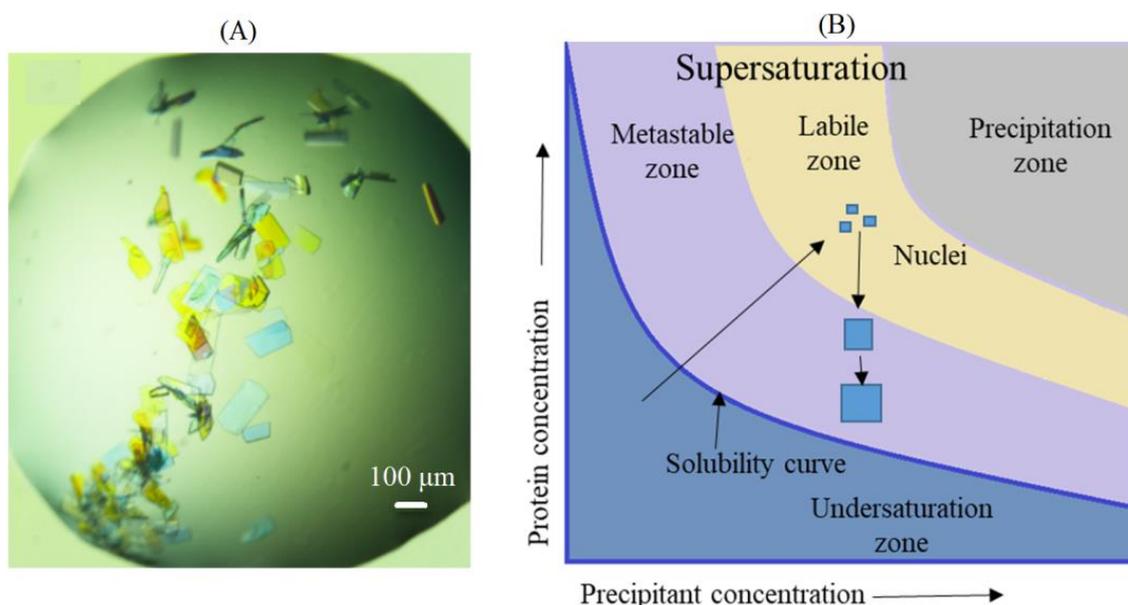


Fig. 3.4 Crystallization of NhGH11. (A) Crystals of NhGH11 were obtained by applying the vapor diffusion hanging drop (B) Phase diagram of crystallization by vapor diffusion. Nuclei formation starts in a labile zone while crystal growth takes place in the metastable zone¹⁶⁶.

The diffusion of water from the crystallization batch increases the protein and precipitate concentration in the drop and as soon as the labile zone is reached, the energy barrier for crystal formation can be overcome to form microscopic protein clusters called nuclei. At this stage, the protein concentration subsequently decreased that creates the metastable zone and crystals start growing^{166,167}. However, microseeding protocols are being used to accelerate the crystallization process and improve the crystal quality. Using micro seeding, the energy barrier for crystal formation has no longer to be overcome and the drop reaches the metastable zone of the phase diagram, in which larger and better quality crystals can grow¹⁶⁸.

3.3.1 Data collection at Synchrotron

High-resolution diffraction data (1.0 Å resolution) was collected at EMBL beamline P13 PETRA III, DESY (Hamburg, Germany). Data collection and refinement statistics are shown in table 3.1. The NhGH11 crystal belongs to the monoclinic space group C2 with unit cell dimensions of $a = 79.45$ Å, $b = 38.50$ Å, $c = 53.59$ Å and $\alpha = \gamma = 90.00^\circ$, $\beta = 91.43^\circ$. Matthews coefficient of NhGH11 crystal was calculated to be 1.9 Å³/Da that corresponds to 37% solvent contents with one NhGH11 molecule in the asymmetric unit.

Table 3.1 Data-collection and refinement statistics

Data collection	
X-ray source	P13 beamline PETRA III, DESY
Detector	Pilatus
Space group	C 2 (No. 5)
Cell dimensions	
a, b, c (Å)	79.45, 38.50, 53.59
α, β, γ (°)	90.00, 91.43, 90.00
Wavelength (Å)	0.97
Resolution range (Å)	34.6 -1.0 (1.03-1.0)
Total number of reflections	258391 (33253)
Redundancy	3.1 (2.6)
Wilson B-factor (Å²)	6.5
R_{meas}	0.053 (0.49)
R_{merge}	0.045 (0.47)
CC_{1/2}	0.999 (0.857)
I/σI	15.57 (2.33)
Completeness (%)	93.12 (78.74)
Refinement	
Reflections used	82543 (6956)
Reflection used for R_{free}	1970 (174)
R_{work}	0.14 (0.24)
R_{free}	0.15 (0.25)
No. atoms	1863
Protein	1580
Water	278
Average B-factor (Å²)	10.78
For macromolecule	8.66
For water	22.61
R.m.s deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.37

Ramachandran	
Favored (%)	97.85
Allowed (%)	2.15
Outliers (%)	0.00
PDB code	6Y0H

*Data in highest resolution shells are shown in parenthesis.

3.3.2 Refinement of single crystal data

The high-resolution structure of NhGH11 was refined by applying the program PHENIX 1.8.4_1496. Initially, default parameters were used for 20 cycles of refinement. Correct amino acids were introduced according to electron density. Further, refinement with isotropic B-values was performed for the full data set and $R_{\text{work}}/R_{\text{free}}$ converged to 0.17/0.18. During the following model building in total 262 solvent water molecules were introduced with occupancies $> 90\%$. The model building was done by applying the program COOT¹⁶⁹.

Electron density was not observed for the N- terminal residues 1 to 4, indicating that these residues were disordered. The high-resolution cutoff of the highest resolution shell (1.03-1.00Å) was determined following the criteria $I/\sigma I \approx 2.3$ and $CC_{1/2} > 30\%$. All atoms except hydrogens were refined anisotropically. Further, 14 additional cycles of anisotropic refinement were performed using all data and including the hydrogen atoms.

The refinement statistics including the reliability factor (R-factor) improved from $R_{\text{work}}/R_{\text{free}} = 0.17/0.18$ to 0.14/0.15 and 102 additional solvent molecules and alternate conformation of 24 residues were introduced. All solvent molecules were introduced with appropriate occupancies and hydrogen bond distances.

Refinement of the NhGH11 structure at atomic resolution allowed the interpretation of static disorders, solvent structure, modeling of atomistic anisotropic vibrations and hydrogen atoms. The average B-factors for all atoms are shown in Fig. 3.5 A. Residues with high B-factors were found on the surface of the NhGH11 structure. The thermal parameters are shown by the ellipsoid view of NhGH11 which indicates that the N-terminus region has a relatively large ellipsoid size as shown in Fig. 3.5 B. These thermal parameters indicate the vibration of atoms in the crystal structure and commonly called atomic displacement parameters (ADPs) or anisotropic temperature factor for high resolution structure. The size of ellipsoids indicates the magnitudes and directions of thermal vibrations of side chains in the atomic resolution structures¹⁷⁰.

High-resolution structures can elaborate structural distortions, especially observed in the bending of aromatic side chains for lysozyme structure measured at 0.65 Å resolution¹⁷¹. However, distorted planarity was also observed for several side chains e.g. twisted guanidinium group or double conformations¹⁷². Comparison of high-resolution crystallographic structures of available similar proteins revealed that such features observed to be conserved within families of proteins, suggesting that these features cannot merely artifacts but true properties of high-resolution crystallographic structures. For example, Laulumaa & Kursula compared less than 1.0 Å resolution structures for PDB entries 1ZUU (0.97 Å), 1TG0 (0.97 Å), 4HVV (0.98 Å), 2G6F (0.92 Å), and 2O9S (0.83 Å) and observed that outlier in aromatic planarity was conserved in all structures, especially close to the binding site and the clearest bending was observed for tyrosine and phenylalanine residues. A crystal structure of myelin protein at sub-atomic resolution was also observed with similar geometrical outliers. The details of unconventional conformations at high resolution were found beyond the standard structure validation algorithms and can provide novel insights into the protein function¹⁷³.

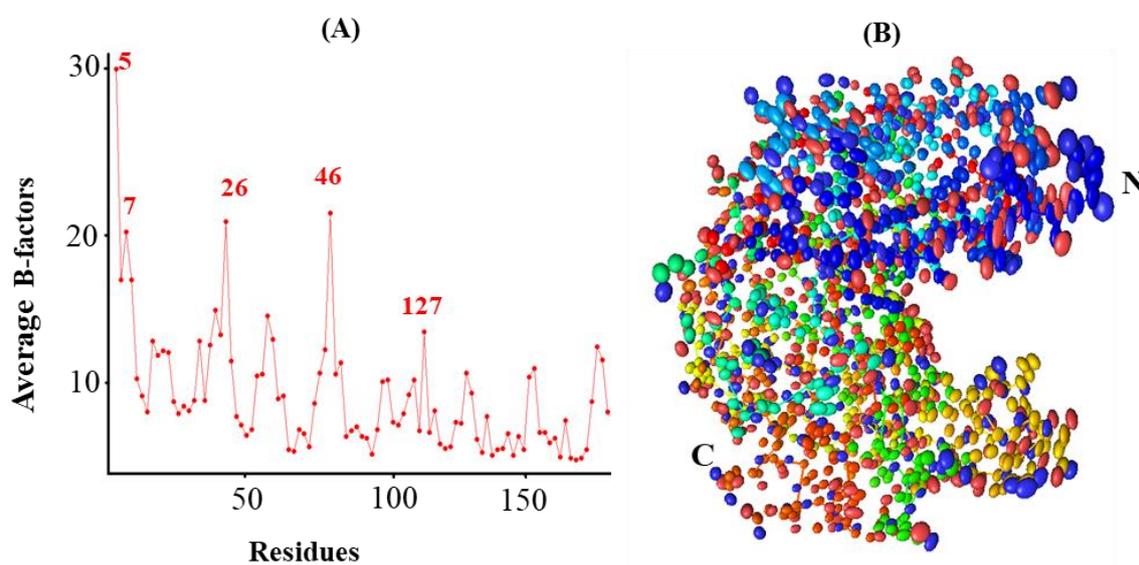


Fig. 3.5 Average B-factors and thermal ellipsoids. (A) Graph indicating peaks for average B-factors for all atoms. (B) The anisotropic thermal ellipsoids representation of NhGH11 in rainbow colors.

The average deviations of chemical bonds and angles between the main chain and side chain atoms were determined for the high-resolution structure of NhGH11. Most of the bond distances and angle restraints were relaxed during refinement cycles. Due to the advantage of high resolution, the deviations from ideal geometries including cis peptide bonds between N56/P57 and N85/P86 were manually inspected.

Some alternative conformations showed deviations from ideal geometries as shown in Fig. 3.6. These conformations were fitted well in the electron density of high-resolution data. Similar unconventional conformations were observed in the atomic resolution (0.72Å) structure of human myelin protein PDB code: 6S2M.

The quality of the NhGH11 structure was analyzed after anisotropic refinement by applying the program PDB REDO¹⁷⁴ and the PDB validation server (<https://validate.rcsb-1.wwpdb.org/>). During refinement, the average errors for chemical bonds and angles were monitored and the structure was deposited in the protein data bank (www.rcsb.org) with PDB code 6Y0H.

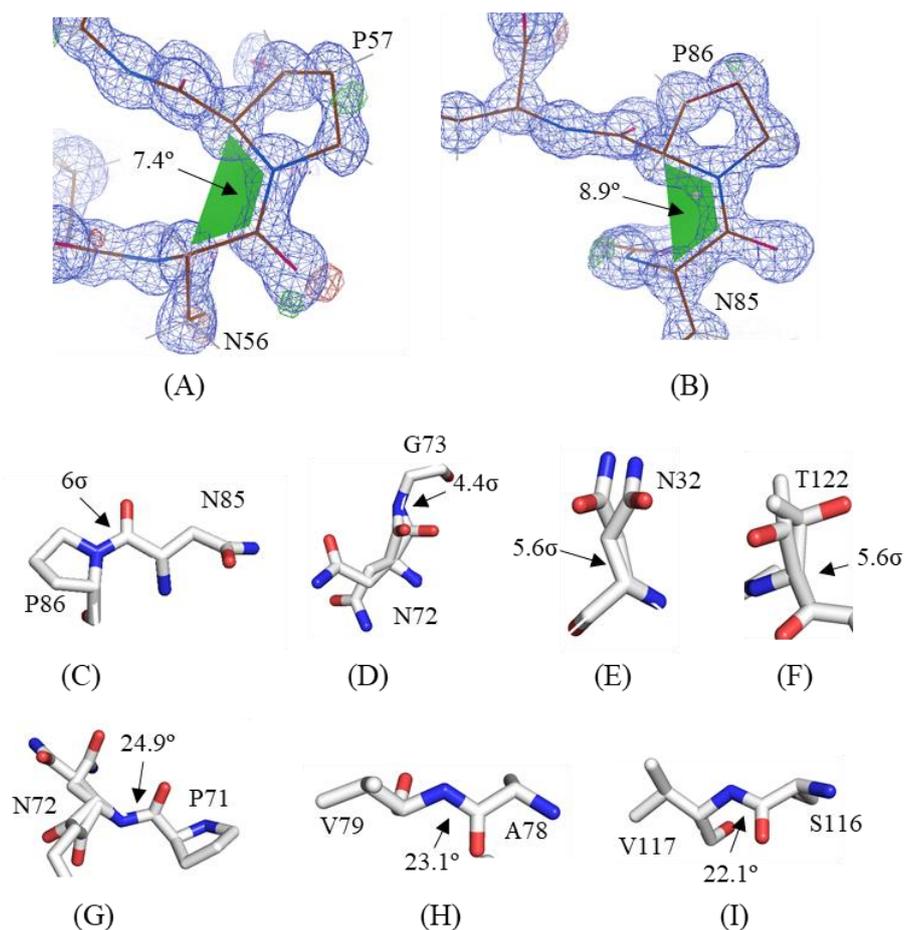


Fig. 3.6 Unconventional conformations. (A) Cis peptide bond between N56/P57. (B) Cis peptide bond between N85/P86. The 2Fo-Fc electron density map is shown at the 1.0 σ level. (C, D) Bond length restraints shown in terms of deviation from ideal values. The bond length between N85/P86 and N72/G73 deviates by 6.0 σ and 4.4 σ from ideal values. For the quantification of dihedral angle restraints, the bending angle distortion is calculated as the deviation from the expected planarity angle 180°. (E, F) Bond angle restraints due to double chain conformations. (G, H, I) The deviation from dihedral angles between P71/N72, A78/V79 and S116/V117 are shown in terms of deviation from ideal geometry. Dihedral angle restraints are calculated as distortion differences from expected planarity (180°).

3.3.3 Crystallographic structure of NhGH11

The crystallographic structure of NhGH11 was solved at high resolution (1.0 Å) which allowed to explore the relationship of this new xylanase with biochemical analysis. NhGH11 provides the first structure of xylanase from *N. haematococca*. The three-dimensional structure of NhGH11 contains the typical β -jelly roll fold of the GH11 family and consists of 193 amino acids. The structure contains two β -sheets with 14 β -strands (Fig. 3.7 A and B). Sheet 'A' consists of five strands and forms a convex-shaped outer layer of β -jelly roll structure; while the inner layer is formed by sheet 'B' which is concave-shaped and contains nine strands. All β -strands are arranged in an antiparallel manner except β 8 and β 11 that are parallel to each other. The structure contains one α -helix that is typically found in all GH11 xylanases. Ten amino acid residues are forming a α -helix, which is located between strands β 12 and β 13. Another 3_{10} -helix is present between strands β 2 and β 3. The three-dimensional structure resembles the right hand with the fingers representing two antiparallel β -sheets and a thumb loop located between strands β 10 and β 11. Approximately, 62% of the total residues are involved in the formation of secondary structure elements. The inner antiparallel sheet is highly twisted to create a well-structured cylindrical cavity where xylan substrates can bind.

The overall structure of NhGH11 presented the characteristics of the family 11 xylanases that allow a detailed comparison with PDB (<https://www.rcsb.org/>) structures of the GH11 family. NhGH11 contains β 1 strand that is absent in many PDB structures like the xylanases from *A. kawachii* (PDB code: 1BK1), *A. niger* (PDB code: 1UKR), *B. circulans* (PDB code: 1XNB), *B. subtilis* (PDB code: 1AXK) *H. jecorina* (PDB code: 1XYN) and *S. acidophilum* (PDB code: 3M4F). Furthermore, the xylanases from *N. patriciarum* (PDB code: 2C1F) and *D. thermophilum* (PDB code: 1F5J) have an additional 3_{10} -helix between β -strands 5 and 6. It is reported that the 3_{10} -helix stabilized the C-terminus of xylanase of *D. thermophilum*¹⁷⁵ via additional hydrogen bonds. Whereas, the 3_{10} -helix in the NhGH11 structure is located between β -strands 2 and 3 where it cannot provide similar interactions with the C-terminus.

The unique, but flexible positions of disulfide bonds were also found in the GH11 xylanases, for example, the xylanases from *A. kawachii* (PDB code: 1BK1), *A. niger* (PDB code: 1UKR) and *P. funiculosum* (PDB code: 1UKR) have disulfide bonds between β 10 and the cord region while the xylanases from *P. varioti* (PDB code: 1PVX) and *T. lanuginosus* (PDB code: 1YNA) have disulfide bonds between β 9 and an α -helix.

While the sequence of NhGH11 does not contain cysteine residues and disulfide bonds are also absent.

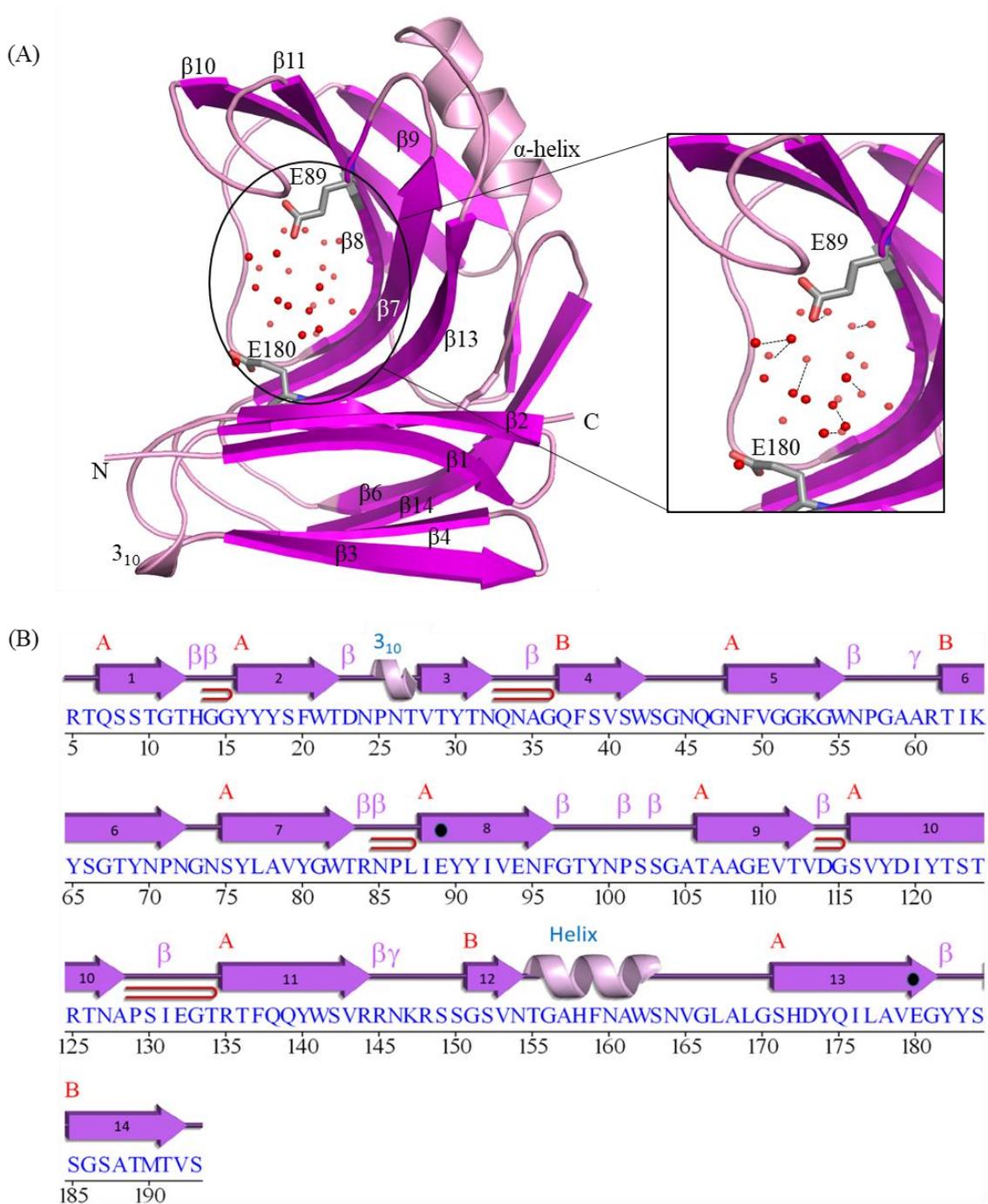


Fig. 3.7 Overall structure of NhGH11. (A) Cartoon plot of NhGH11 with active site residues, shown in sticks. A zoom of the active site region is shown (right) with catalytic residues surrounded by solvent molecules in the active site cavity. (B) Sequence and secondary structure with 14 strands (purple), helices (pink) and other motifs (in red β -hairpins, A and B sheets). Catalytic residues are indicated by black circles. The figure was prepared by applying the PDBsum server (<http://www.ebi.ac.uk/pdbsum>).

3.3.4 Surface exposed aromatic amino acids

The surface of NhGH11 was observed to be rich in aromatic amino acids. The surface exposed aromatic amino acids i.e., Y15, Y16, Y17, W21, W54, Y64, Y68, F158 and W161 can play an important role in the substrate selectivity of NhGH11. To understand the role of these aromatic amino acids, the three-dimensional structure of NhGH11 was compared with *B. subtilis* xylanase (PDB code: 1BCX). NhGH11 showed 55% sequence similarity to 1BCX with a root mean square deviation (RMSD) of 0.85 Å for the Ca atoms. The surface of 1BCX displayed aromatic amino acids i.e., F48, Y94, Y113 and W185, which were not spatially conserved with NhGH11 residues (Fig. 3.8). The role of surface exposed residues of 1BCX in substrate selectivity has been studied, and it has been reported that the mutations of F48, Y94, Y113 and W185 decreased the substrate selectivity and the most drastic effect was observed for W185¹⁷⁶. The comparison of both structures showed that the site-directed mutagenesis of surface exposed residues in combination with activity assays can provide a complete understanding of surface exposed residues and their functional role in substrate selectivity.

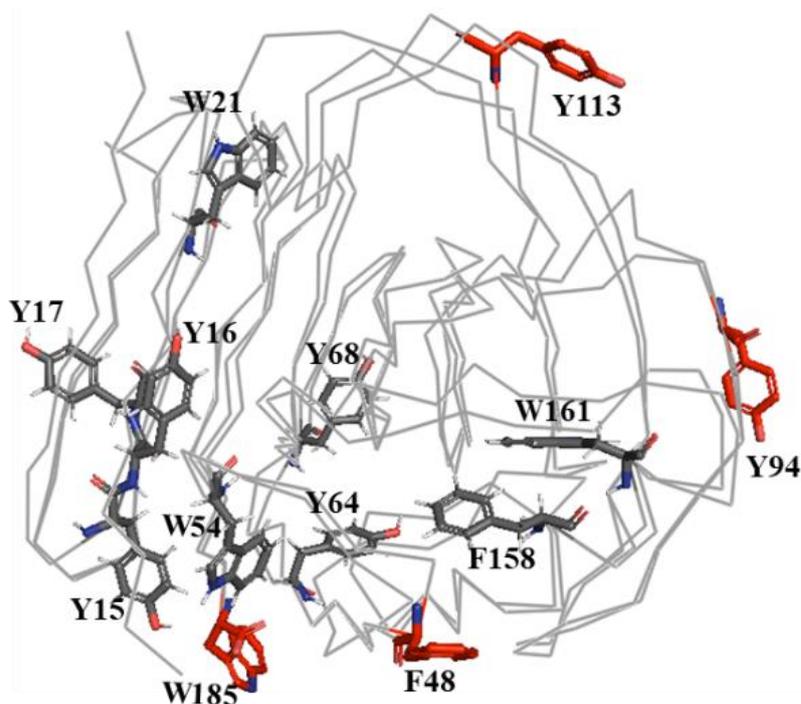


Fig. 3.8 Surface exposed aromatic amino acids of NhGH11. Superimposition of NhGH11 and 1BCX along with the presentation of aromatic amino acids on the surface. The aromatic residues are shown in red color (1BCX) and gray color (NhGH11). Residue numbers show the positions of residues in the respective sequence.

3.3.5 Catalytic cavity of NhGH11

The substrate-binding cavity of NhGH11 contains catalytic residues E89 and E180 that are located with a side chain distance of ~ 10.5 Å. The cavity is ~ 11 Å deep, ~ 4.5 Å wide and ~ 26 Å long. The active site residues are surrounded by several aromatic amino acids i.e. W21, Y76, Y80, Y91, Y174 and Y182 (Fig. 3.9). The overall structure of NhGH11 and active site are homologous to the xylanase from *T. reesei* (PDB code: 4HK8) with an RMSD value of 0.51 Å for Ca atoms. The roles of NhGH11 active site residues were proposed after superimposing the structure on 4HK8 that is complexed with xylohexose. NhGH11 can process a similar substrate and the comparative analysis confirmed that the extended active site cleft of NhGH11 can also accommodate a linear xylan backbone of six xylose sugars.

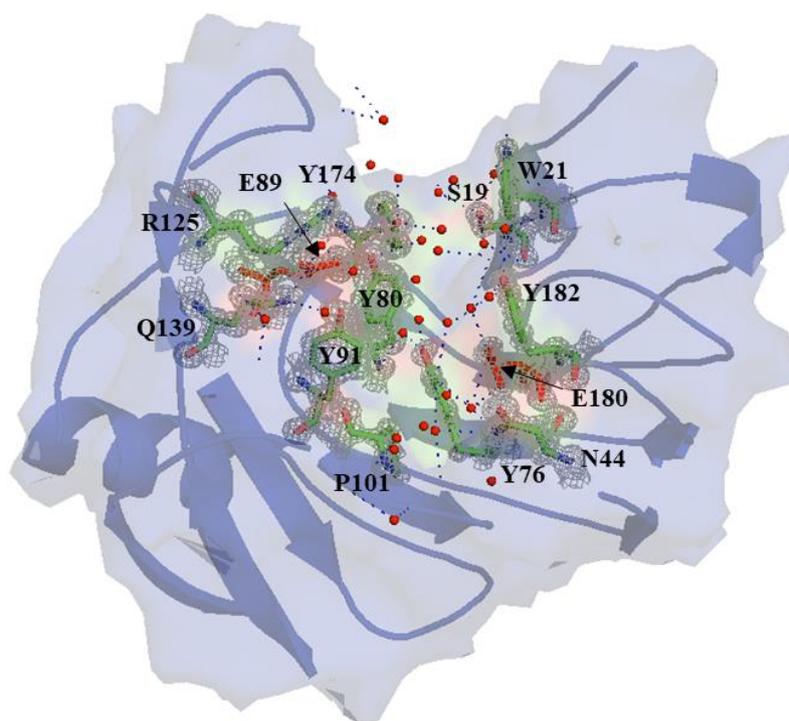


Fig. 3.9 Active site cleft of NhGH11. Cartoon representation with electron density for residues lining the active site cleft and forming hydrogen bonds with solvent molecules. Catalytic residues are indicated with black arrows.

Inside the active site of NhGH11, the carboxyl group of E89 (OE1) forms a hydrogen bond to nitrogen (NE2) of Q139, hydrogen (HG3) of F137 and a water molecule (W49), while the second oxygen of the carboxyl group (OE2) is forming hydrogen to Y80 (Fig 3.10 A). The second catalytic residue, E180 is forming hydrogen bonds with two water molecules via oxygen atoms OE1 and OE2 of the carboxyl group. E180 is strongly bent

towards the main chain because the oxygen atom OE1 was forming a hydrogen bond with the main chain nitrogen of S75. The details of hydrogen bonding and alternative conformations of amino acid side chains are visible in the high-resolution structure of NhGH11. The positions of hydrogen atoms provided information about the complex hydrogen bond network in the active site. Several very clear positions of hydrogen atoms can be observed in high resolution structure, two of them are indicated by black arrows in Fig. 3.10 A. Also, the solvent molecules with alternative conformations are indicated in Fig. 3.10 B.

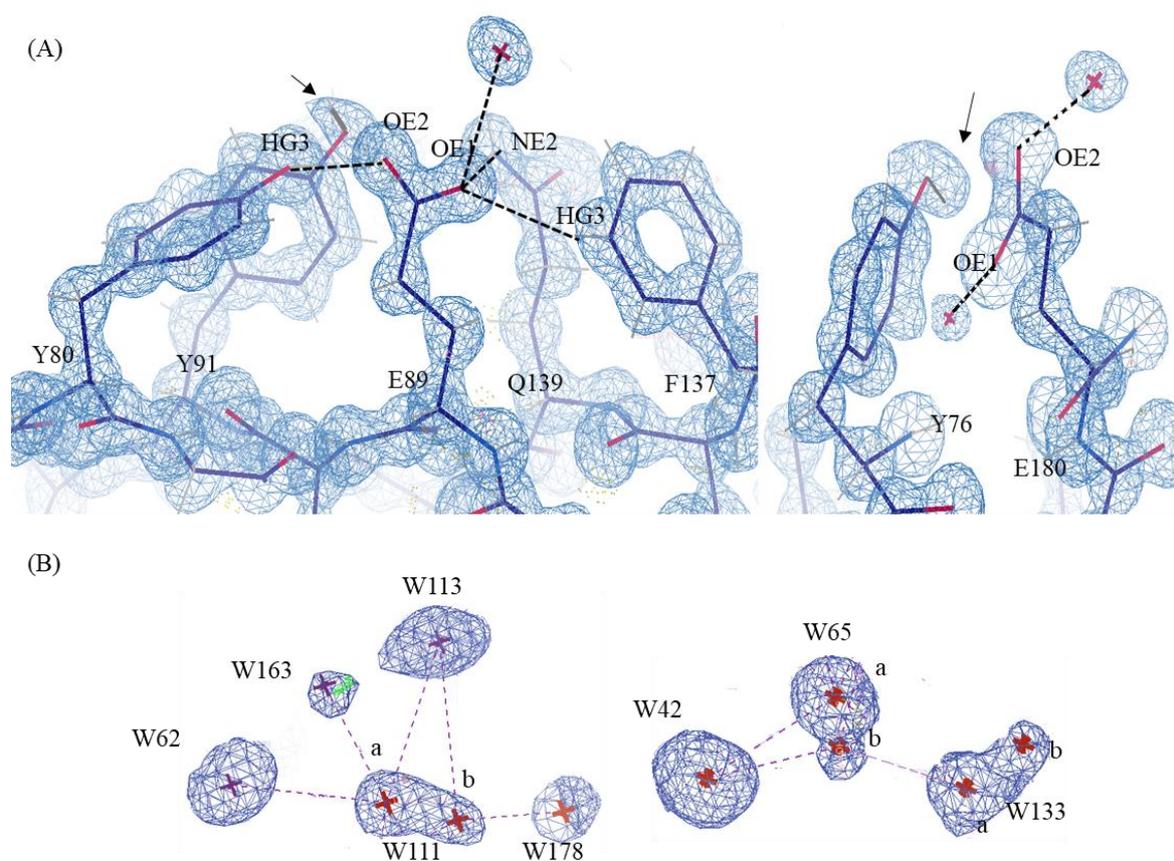


Fig. 3.10 Insights into hydrogen bonding in the active site cleft. (A) Catalytic residues E89 and E180 are shown with vicinal residues. Hydrogen atoms are visible for Y76 and Y91, indicated with arrows. (B) Solvent structure with alternative conformations, “a” conformations of W65, W133 and W111 with corresponding refined occupancies of 57%, 66% and 55% while “b” conformations showed 43%, 34% and 45% occupancies respectively. 2Fo-Fc electron density maps are shown at 1.0 σ level. All hydrogen bonds have distances of ~ 3.0 Å.

3.3.6 Analysis of clefts and binding sites for sugars

Analysis of surface clefts was carried out by the PDBsum server. The results showed that NhGH11 has some potential binding clefts on the surface (Fig. 3.11). The role of these surface binding clefts is not fully understood but xylanase from *P. simplicissimum* (PDB code: 1B3V) provided evidence of xylose binding on the surface. Several other enzymes involved in carbohydrate metabolism have sugar binding clefts¹⁷⁷.

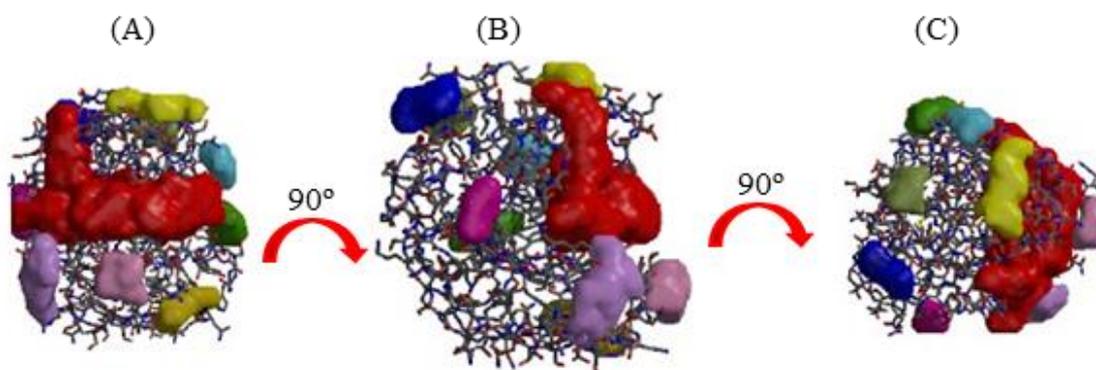


Fig. 3.11 Representation of binding clefts on the surface of NhGH11. Potential binding clefts are shown with solid surfaces. Different colors are highlighting the volume of clefts and the largest cleft regions are shown in red color. The clefts of NhGH11 were analyzed by the PDBsum server¹⁷⁸.

Bioinformatic analysis of substrate binding sites of NhGH11: COACH server (<https://zhanglab.ccmb.med.umich.edu/COACH/>) was applied for prediction of ligand binding sites that provided the complementary binding site by using comparative methods¹⁷⁹. These predictions provide the combined results from protein interaction servers named COFACTOR, FINDSITE and ConCavity to generate final ligand binding site predictions as shown in table 3.2.

Table 3.2 Prediction of consensus binding site residues of GH11 xylanase

PDB code	C-score	Ligands	Consensus binding residues
COACH results			
1H4G	0.77	1, 2-Deoxy-2-Fluoro-4-o- β -D-xylopyranosyl- β -D-xylopyranose	15, 17, 46, 48, 76, 78, 85, 87, 121, 125, 126, 133, 135
3WP6	0.58	β -D-Xylobiopyranose	15, 17, 46, 48, 76, 78, 125, 126, 127
4HKW	0.18	β -D-Xylopyranose	44, 72, 87, 121, 135, 137, 176, 178
1REF	0.11	2, 3-Epoxypropyl- β -D-Xyloside	76, 78, 85, 87, 126, 170, 176

2VGD	0.06	β -D-Xylopyranose	70, 72, 95, 178
TM-SITE results			
1C5I	0.39	β -D-xylopyranose-(1-4)-1, 5-anhydro-2-deoxy-2-fluoro-D-xylitol	17, 44, 46, 76, 78, 85, 87, 121, 125, 133, 135, 176
2VGD	0.17	β -D-xylopyranose-(1-4)- β -D-xylopyranose-(1-4)- β -D-xylopyranose	17, 126, 127
COFACTOR Results			
1RED	0.74	4, 5-Epoxyptentyl- β -D-xyloside	17, 46, 76, 85, 87, 126, 170
4HK8	0.71	Xylohexose	17, 44, 46, 76, 78, 85, 87, 121, 125, 135, 176

*C-score is indicating the confidence score of the prediction results. C-score ranges from 0-1; a higher score indicates the high reliability of prediction.

Docking: xylohexose was docked to the active site cavity of NhGH11 by applying the Flexidock program, which belongs to the Sybyl-X program package (Tripos, USA). Docking results were compared with xylohexose bound structure of xylanases PDB code: 4HK8. Six subsites were identified in the NhGH11 structure, the residues W21, S19, Y76, Y80, E89, Y91, P101, E180, R125, Q139, Y174 and Y182 were found in active site cleft that can interact with the substrate (Fig. 3.12).

The possible interactions at the binding site are explained by a combination of superimposition (PDB code: 4HK8) and docking calculations:

Subsite -3: the binding of the xylose subunit at subsite -3 was stabilized by W18 by a water-mediated hydrogen bond in the structure 4HK8. It is speculated that W21 can also provide the same interactions in NhGH11.

Subsite -2: aromatic amino acids occupied this subsite to provide the stacking interactions with the xylose subunits of the substrate²⁸. In PDB structure 4HK8, the hydroxyl groups of Y171 and Y76 are providing direct hydrogen bonding and S16 was providing water-mediated hydrogen bonding to the hydroxyl group of xylose sugar. In NhGH11, the subsite -2 displayed the same residues as for 4HK8 that can provide the same interactions via hydroxyl groups of Y174, Y80, and S19.

Subsite -1: this site displays the nucleophilic catalytic residue E86 and R122, which provided direct hydrogen bonds to the xylose subunit in the 4HK8 structure. The same interaction can be observed by E89 and R125 in the active site of NhGH11.

Subsite +1: catalytic acid/base residue E177 provides direct hydrogen bonding to an oxygen atom of the glycosidic bond between xylose subunits at -1 and +1 for 4HK8. These interactions cannot be observed for NhGH11 due to the bent side chain of E180 previously shown in Fig. 3.10 A.

Subsite +2: a residue R122 provides a water-mediated hydrogen bond and N71 provides a direct bond to a xylose subunit for 4HK8. In the NhGH11 structure, R125 and N74 can provide the same interactions.

Subsite +3: a residue N71 provides a direct hydrogen bond to the substrate while Y179 provides a water-mediated hydrogen bond. In the NhGH11 structure, N74 and Y182 can provide similar interactions.

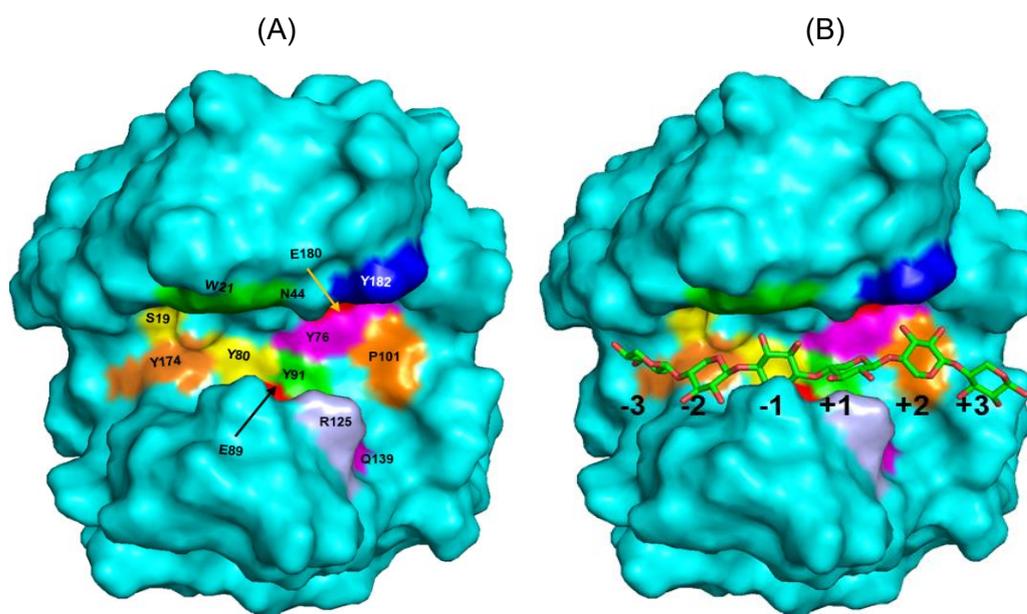


Fig. 3.12 Docking of xylohexose to the substrate-binding cleft of NhGH11. (A) The substrate-binding cleft is indicated in different colors. The xylohexose was docked to the substrate-binding cleft. Possible interacting residues in the cleft are shown in different colors. (B) The active site cleft of NhGH11, showing a binding site with six subsites.

3.4 Sequence homology

A detailed comparison of GH11 sequences was carried out for 82 sequences, from which 46 of the selected sequences belong to fungi and 36 to bacteria. Sequences of the GH11 family were selected from the carbohydrate-active enzyme database (<http://www.cazy.org/>). The FASTA sequences were obtained from the universal protein resource (UniProt: <https://www.uniprot.org/>) and alignment was performed by server Clustal omega (<https://www.ebi.ac.uk/Tools/msa/clustalo>). The comparison of secondary

structure and conserved regions of NhGH11 was performed by server Esript 3 (<http://esript.ibcp.fr/ESript/ESript/>) as shown in the supplementary information, S. 4.2. The highest sequence homology was observed for the β -strands $\beta 5$, $\beta 8$, $\beta 10$ and $\beta 11$ also the C-terminal region (strands 12 to 14) was highly conserved.

An evolutionary tree was constructed for the same 82 sequences by applying the program Geneious prime (<https://www.geneious.com/>) shown in the supplementary information, S. 4.3. The results showed that NhGH11 shared common ancestors with fungal GH11 xylanases and more closely related to a common ancestor with endo- β -1,4-xylanase from *Magnaporthe grisea* (UniProt ID: Q92244).

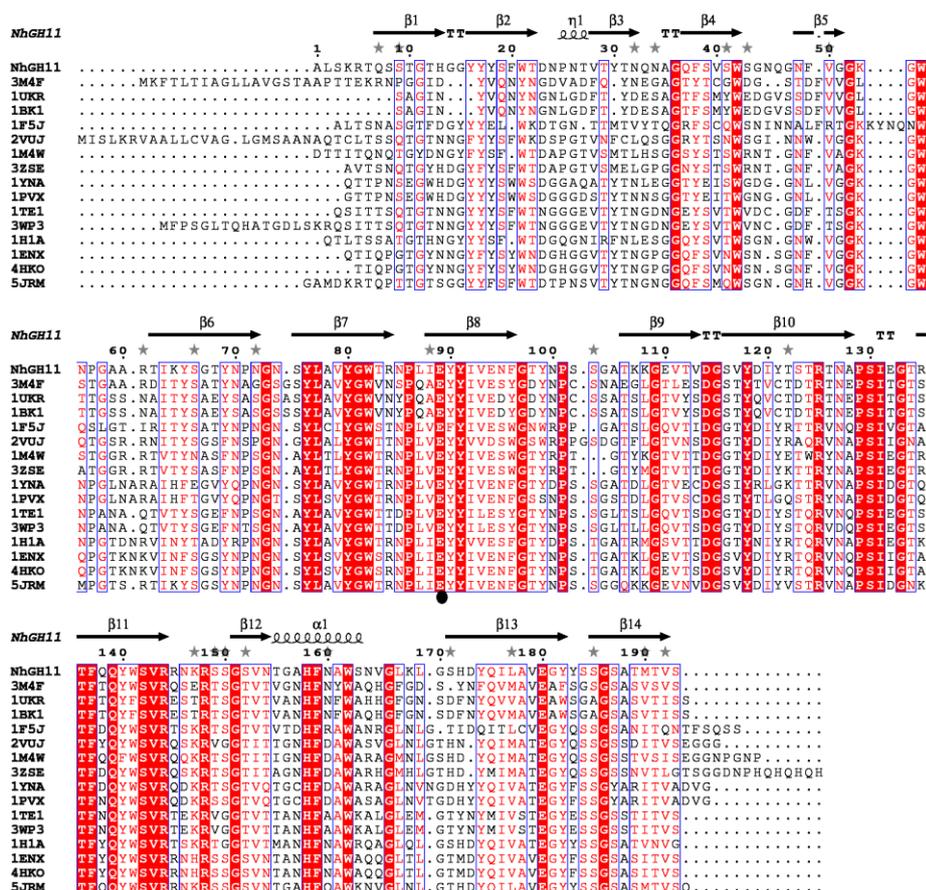
To identify the most conserved regions, the 3D structure of NhGH11 was compared to the GH11 PDB structures i.e., 5JRM, 4HK0, 1ENX, 1H1A, 1PVX, 1YNA, 1M4W, 3ZSE, 1TE1, 3WP3, 2VUJ, 1F5J, 1UKR, 1BK1 and 3M4F, from *F. oxysporum*, *T. reesei*, *T. reesei*, *C. thermophilum*, *B. spectabilis*, *T. lanuginosus*, *T. flexuosa*, *T. fusca*, *P. funiculosus*, *T. Funiculosus*, *E. coli*, *D. thermophilum*, *A. niger*, *A. luchuensis* and *K. pastoris*, respectively. Aligned PDB structures showed 82%, 71%, 71%, 66%, 62%, 62%, 60%, 59%, 59%, 56%, 52%, 51%, 48%, 48% and 47% sequence similarities respectively. Secondary structure regions indicated with gray stars showed the residues with alternative side-chain conformations in the high-resolution structure of NhGH11 (Fig. 3.13 A). The three-dimensional structure of NhGH11 superimposed with the homologous structure is shown with the backbones in red color indicating the regions with sequence identity for all structures (Fig. 3.13 B).

Superimposition results of endo β -1,4 xylanase PDB structures showed a root mean square deviations (RMSD) for all residues less than 1.5 Å for the C α atoms that correlate well with the observed sequence similarities, highlighted and shown for spatially conserved regions. Approx. 60% of the secondary structure elements correspond to β -strands and ~5% to a α -helix. The thumb-like loop between $\beta 10$ and $\beta 11$ strands showed highly conserved amino acids P129, S130, and I131 that were found to be important for enzyme activity and releasing of the products after catalysis, therefore, deletion of these residues drastically changed the active enzyme conformation to an inactive form²⁸. Aromatic residues W42, W55, Y69, Y76, Y80, W82, Y90, Y91, F96, Y118, F137, W141, F158, W161, Y175 and Y182 are highly conserved, while F21, W22 and Y66 are less conserved. These aromatic amino acids are important for substrate binding, e.g. *B. circulans* xylanase the mutagenesis of Y77 and Y88 to phenylalanine resulted in complete inactivation of

3. Results and discussion

enzyme¹⁸⁰. Similar effects were reported for mutation of Y73 and Y88 to alanine in *B. subtilis* xylanase¹⁸¹.

(A)



(B)

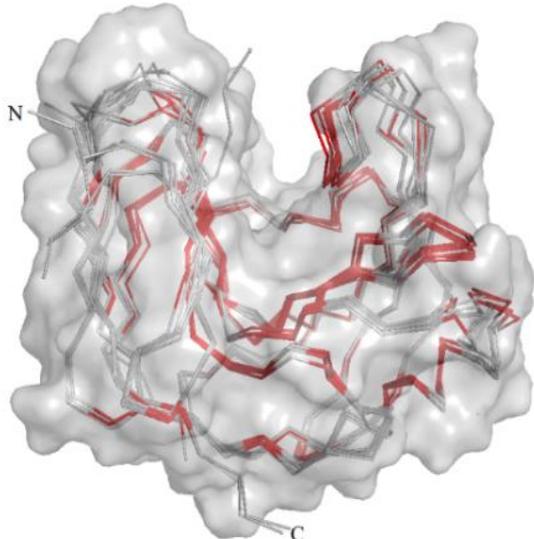


Fig. 3.13 Sequence alignment and 3D comparison. (A) Residues in red color are identical and residues in pink color show similarity among aligned sequences. Catalytic residues E89 and E180 are indicated with filled circles (B) Surface representation with underlaid Ca tracing, red color indicates regions with sequence identity for all structures.

The strictly identical and spatially conserved active site residues from PDB structures of GH11 xylanases were superimposed with NhGH11 residues (Fig. 3.14) and the RMSD values of individual residues (Table 3.3) were calculated by applying a program UCSF Chimera¹⁸². Spatially conserved active site residues indicate the highly conserved substrate binding residues for GH11 family.

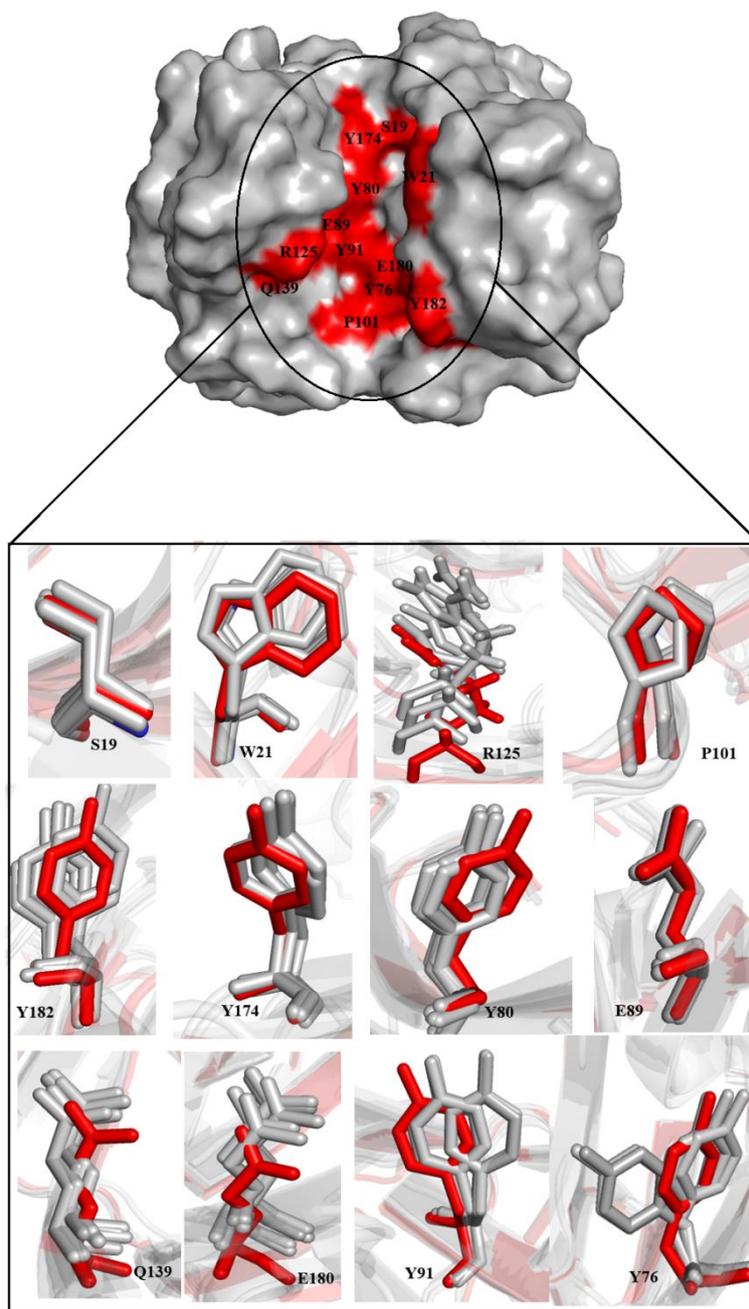


Fig. 3.14 Surface representation and comparison of spatially conserved active site residues. NhGH11 was superimposed with homologous structures (PDB codes: 5JRM, 2JIC, 4HK8, 1H1A, 1YNA and 1PVX) from *F. oxysporum*, *T. longibrachiatum*, *H. jecorina*, *C. thermophilum*, *T. lanuginosus* and *B. spectabilis* respectively. NhGH11 residues are shown in red while the other six homologous residues are indicated in grey color. Residue numbers are assigned referring to NhGH11.

Table 3.3 The root mean square deviation of individual residues (Å) in the active site cleft

	5JRM	4HK8	2JIC	1H1A	1YNA	1PVX
S19	0.43	0.46	0.55	0.79	0.68	0.55
W21	0.51	0.96	0.87	1.50	0.63	0.34
Y76	0.65	0.68	0.43	0.68	0.36	0.60
Y80	0.41	0.61	0.54	0.54	0.47	0.52
E89	0.80	1.23	0.55	1.23	0.64	0.51
Y91	0.36	0.60	0.47	0.51	0.26	0.25
P101	0.18	1.15	0.68	2.20	0.41	0.42
R125	1.95	2.82	1.52	2.80	1.06	1.00
Q139	1.43	1.64	1.00	1.78	1.15	0.88
Y174	0.22	0.40	0.54	0.27	0.58	0.65
E180	0.43	0.60	0.72	0.69	0.28	0.32
Y182	1.00	1.17	0.97	1.09	0.73	0.84

*The highest differences between residues of NhGH11 and other structures are indicated in bold.

3.5 Elucidation of thermostability parameters

The thermostability was analyzed in terms of thermal unfolding of the secondary structure of NhGH11 using circular dichroism (CD) spectroscopy. The results indicated that the secondary structure of the NhGH11 was stable in an aqueous solution up to 45 °C. Above 45 °C, the α -helical contents started decreasing, but β -sheets remained partially stable up to 90 °C (Fig. 3.15). The results of thermal unfolding experiments showed that NhGH11 is highly stable up to 45 °C (Fig. 3.15, red curve).

Extensive studies have been performed to identify structural features of xylanases related to thermostability¹⁸³. Structural parameters supporting the thermostability of xylanases included the number of disulfide bonds, hydrogen bonds and aromatic side chains forming stacking pairs and buried water molecules¹⁸⁴. The thermostability parameters of NhGH11 were compared with GH11 xylanases from *D. thermophilum*¹⁸⁵, *C. thermophilum*⁴¹, *T. lanuginosus*³¹, *P. varioti*³³ and *T. reesei*⁴⁰. Interactions and sequence factors contributing to the mesophilic character of NhGH11 in comparison to thermophilic xylanases are summarized in table 3.4.

All thermophilic xylanases showed a higher number of side-chain to side-chain hydrogen bonds as compared to NhGH11 and thermophilic xylanases from *C. thermophilum*, *T. lanuginosus* and *P. varioti bainier* showed more ion pairs compared to NhGH11. The comparison of twelve xylanases concerning their thermostability parameters indicated some minor modifications such as higher Arg/Lys and Thr/Ser ratio, or lower number of Asn and Gln residues that could enhance the thermostability⁴⁰.

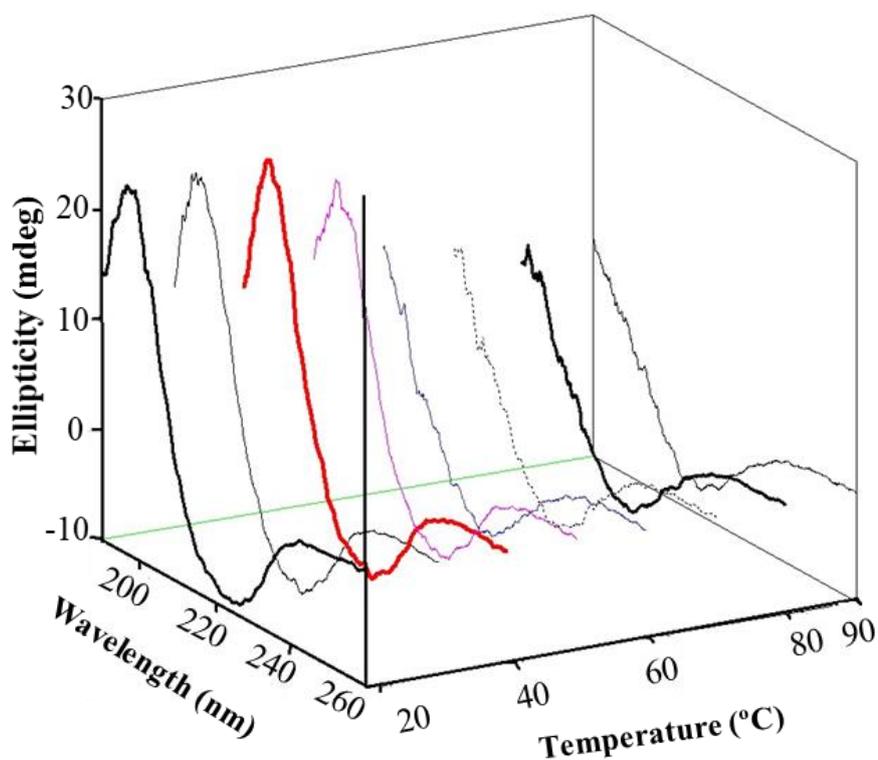


Fig. 3.15 Thermostability in terms of thermal unfolding of secondary structure. Spectra by Circular Dichroism (CD) are shown as a function of temperature from +20 to +90 °C, each measurement of thermal unfolding was performed by 10 °C intervals.

Most of the thermophilic xylanases shown in table 3.4 have a higher number of Arg/Lys and Thr/Ser ratio as compared to NhGH11. Thermophilic xylanases from *P. varioti bainier*, *T. lanuginosus* and *C. thermophilum* had a lower number of Asn and Gln residues as compared to NhGH11.

Another structural feature supporting thermostability of *D. thermophilum* xylanase is a 3_{10} -helix located in the middle of the structure and also stabilizing the C-terminus, which is not present in NhGH11. This 3_{10} -helix provided 34 hydrogen bonds¹⁷⁵, compared to 14 hydrogen bonds formed by the corresponding region of NhGH11 structure.

A calculation of solvent accessibility showed that NhGH11 contained seven buried water molecules (Fig. 3.16 A). Aromatic side-chain forming π - π and C-H- π interactions included the residue side chains of F20-Y30-F38, W42-F49, Y91-W141, W82-F137 and Y182-Y183 can also contribute to the overall thermostability of NhGH11 (Fig. 3.16 B). Aromatic amino acid pairs are supposed to contribute 1-2 kcal mol⁻¹ in terms of thermostability^{186,187}.

Table 3.4 Thermostability factors of NhGH11 in comparison to thermophilic xylanases

Xylanase	NhGH-11	<i>D. thermo-philum</i>	<i>C. thermo-philum</i>	<i>T. lanugi-nosus</i>	<i>P. varioti</i>	<i>T. reesei</i>
PDB code	6Y0H	1F5J	1H1A	1YNA	1PVX	1XYN
Number of residues	189	199	191	194	194	178
Optimum temperature (°C)	45	75	60	65	65	50
Ion pairs	10	10	18	18	13	9
Hydrogen bonds (main-chain-main-chain)	163	163	163	163	157	152
Hydrogen bonds (main-chain-side-chain)	77	71	70	67	69	55
Hydrogen bonds (side-chain-side-chain)	61	89	71	81	77	68
Aromatic interactions	11	15	12	20	19	13
Hydrophobic interactions	140	160	144	127	115	135
No of Pi-Pi interactions	6	7	6	6	6	6
Arg/Lys ratio	2.6	2.5	3.3	2.6	3.0	3.0
Thr/Ser ratio	0.87	1.40	1.41	1.38	0.95	0.78
Asn+Gln	26	28	24	20	21	29
Cys residues	0	3	0	2	2	0
Aromatic residues	32	29	30	30	30	22
Buried waters	7	7	6	6	7	8

*The highest differences between thermophilic GH11 enzymes and mesophilic NhGH11 are indicated in bold.

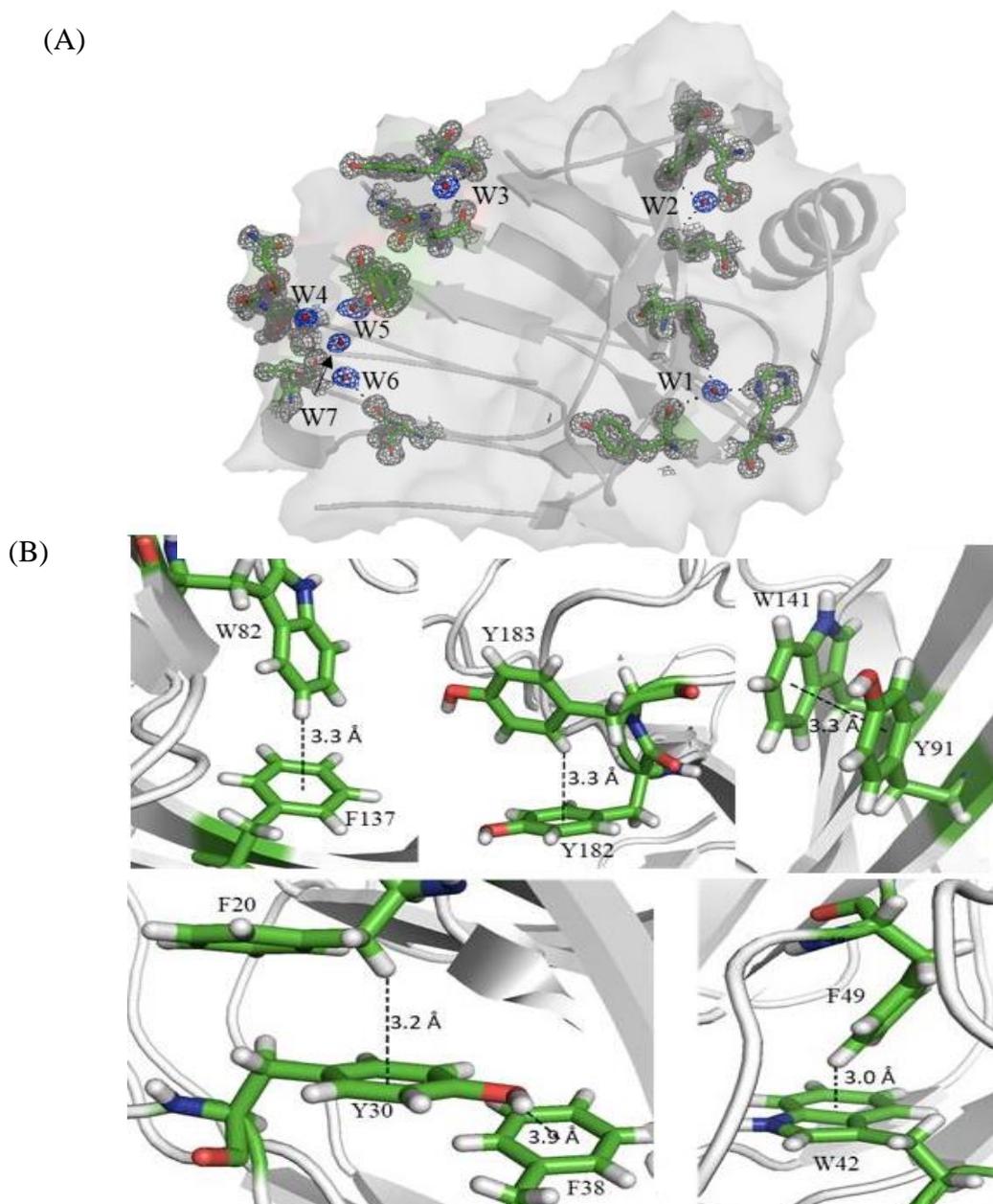


Fig. 3.16 Thermostability parameters (A) Cartoon representation showing buried water molecules (W1-W7) with electron density within the cartoon surface representation. $2F_o - F_c$ electron density maps for water molecules and surrounding residues are shown at 1.0σ level. (B) Selected intramolecular aromatic-aromatic (π - π) interactions and C-H- π interactions. These interactions are considered if the distance between the interacting proton and centroid of the phenyl group is less than 3.5 \AA .

The temperature-dependent stability of NhGH11 was analyzed at different temperatures ($40 \text{ }^\circ\text{C}$, $45 \text{ }^\circ\text{C}$ and $50 \text{ }^\circ\text{C}$) for 8 hours as shown in Fig. 3.17. Activity at optimum conditions was considered 100%. After 4 hours, NhGH11 showed 80% activity at $40 \text{ }^\circ\text{C}$ and $45 \text{ }^\circ\text{C}$ while the activity decreased to approx. 50% after 8 hours. At $50 \text{ }^\circ\text{C}$, the enzyme lost approx. 45% of its activity after 4 hours that further decreased to 32% at eight hours.

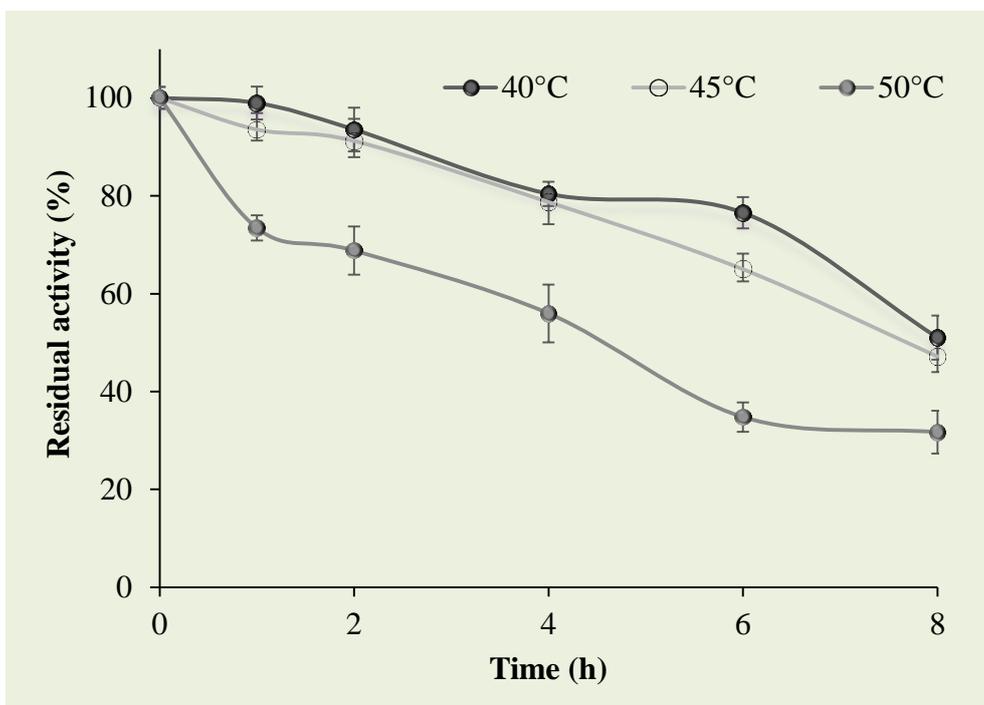


Fig. 3.17 The temperature-dependent stability and activity. The residual activity was evaluated at 40 °C, 45 °C and 50 °C for 8 hours. The data points represent the mean values of three replicates and error bars to indicate the corresponding standard deviation.

The optimum temperatures for most of the xylanases are in a range between 35 °C to 85 °C and the temperature preference categorized the xylanases as mesophilic or thermophilic enzymes²⁸. The xylanase from *T. lanuginosus* (optimum temperature 65 °C) showed 64% residual activity after 8 h incubation at 60 °C and xylanase from *T. fusca* (optimum temperature 75 °C) showed 96% residual activity after 18 h of incubation at 75 °C²⁸. The thermostability of GH11 xylanases evolved from the following structural differences of GH11 xylanases.

Disulfide bonds: strengthen the overall framework and enhance the stability of xylanases. The introduction of disulfide bonds in *T. resei* xylanase enhanced the thermostability to 2.5 folds as compared to wild type¹⁸⁸.

Ion pairs and aromatic pairs: a higher number of ion pairs and aromatic-aromatic interactions favor the formation of hydrogen bond interactions and provide a more stable structure by the formation of intramolecular interactions⁴⁰.

Oligomerization: oligomerization of xylanases also enhance the thermostability, the effect was observed for the thermostability of *T. xylanilyticus* xylanase¹⁸⁹. Besides, the presence of aromatic residues on the surface of xylanase creates hydrophobic patches that can also

trigger the dimerization of monomers and enhance the overall thermostability⁴⁰.

Overall thermostability originates from a combination of different factors and thus none of the individual features can be considered as absolute alone. Thermostability factors can be considered as the individual property of a specific enzyme. The structure based improvements of the thermostability of xylanases are ongoing, due to high temperature requirement of industrial processes¹⁹⁰. However, the thermostability mutagenesis studies are complex and till now not well understood. In this context, the first mutagenesis study was carried out for a *B. pumilus* xylanase and a double mutant G38S/R48K was created which showed a shift of half-life from 4 to 12 min at 57 °C¹⁹¹. Another investigation was done with *Streptomyces* sp. S38 e.g. the mutations of hydrophobic residues located within β -sheets enhanced the optimum temperature from 57 to 66 °C¹⁹². Also, the incorporation of the N-terminus of thermostable xylanase into mesophilic xylanase improved the thermostability. For example, mesophilic xylanase displayed an optimum temperature of 55 °C and a half-life of few minutes at 70 °C, but integrating the N-terminus of thermophilic xylanase improved the optimum temperature and enhanced the half-life time up to 120 minutes¹⁹³. The optimum temperature of a mesophilic *S. olivaceovirdis* xylanase was also improved by 15 °C when the N-terminus of thermophilic xylanase was incorporated¹⁹⁴.

Besides mutagenesis studies, the disulfide bonds also improved the thermostability e.g. the introduction of two disulfide bonds between the N- and C-terminus and between strand B9 and α -helix lead to enhanced thermostability¹⁹⁵. In another study creation of a disulfide bond between the N-terminus and the adjacent strand resulted in the enhancement of the temperature optimum from 57 °C to 70 °C and also half-life time was increased¹⁹⁶. GH11 xylanases contain one highly conserved α -helix that seems sensitive to thermal unfolding by applying temperatures higher than optimum values. Therefore, the thermostability of GH11 xylanases can be improved by making this α -helix region more stable¹⁷⁵.

3.6 Tape drive experiments

The crystallographic studies via the tape drive mix-and-diffuse serial crystallography were performed at P11 synchrotron beamline PETRA III, on DESY campus which allowed the data collection of NhGH11 in presence of xylopentose at mixing times of few seconds. The tape drive experiments were performed by microcrystals suspension of NhGH11 with crystals dimensions between 10 to 20 μm .

Crystals were obtained from two different precipitant conditions:

Precipitant 1: 1M $(\text{NH}_4)_2\text{SO}_4$, 100 mM sodium citrate pH 5.5.

Precipitant 2: 200 mM $(\text{NH}_4)_2\text{SO}_4$, 100 mM sodium citrate pH 5.5 20% PEG 6000.

Crystals from both conditions were applied for data collection. Microcrystals grown under the precipitant 1 condition (Fig. 3.18 A) showed 50:50 probabilities for two space groups P2 and C2, therefore, further data processing was not possible. On the other hand, microcrystals are grown in precipitant 2 (Fig. 3.18 B) presented good quality data with the maximum probability for space group C2. The tape drive experiments were set up by the collaboration partner Dominik Oberthür. Before data collection, the polyimide tape was fixed between two rollers under tension and a needle for sample delivery was fixed so that X-rays pass through a hole in the center of the device, the experimental setup is shown in Fig. 3.18 C. The needle for sample delivery had special mixer conformation for mixing of microcrystals and substrate as shown in Fig. 3.18 D. Xylopentose (70 mM) was dissolved in a precipitant 2 to be applied for mixing times of 1, 2, 3, 5 and 7 seconds.

Dominik Oberthür integrated and merged reflection files of data sets by applying a program CrystFEL version 0.6.1¹⁹⁷.

The refinement of the initial electron density maps obtained at different mixing was performed by using the NhGH11 model obtained by single crystal diffraction. Unfortunately, the inspection of the electron density maps at different σ levels did not provide any evidence of substrate binding. The crystal packing of NhGH11 was probably too tight resulting in the surface molecules hindered the penetration of the substrate into the crystals. The details of crystal packing are provided in section 3.7. However, the superimposition of the high-resolution structure from cryogenic conditions to the room temperature structure of NhGH11 (Fig. 3.18 E) showed a non-significant RMSD (0.11Å) for C α atoms. Also, the comparison of data collection table 3.1 and table 3.5 showed that both data sets have the same space groups C2 while the obtained resolution of the data at room temperature is lower than the data collected under cryogenic conditions.

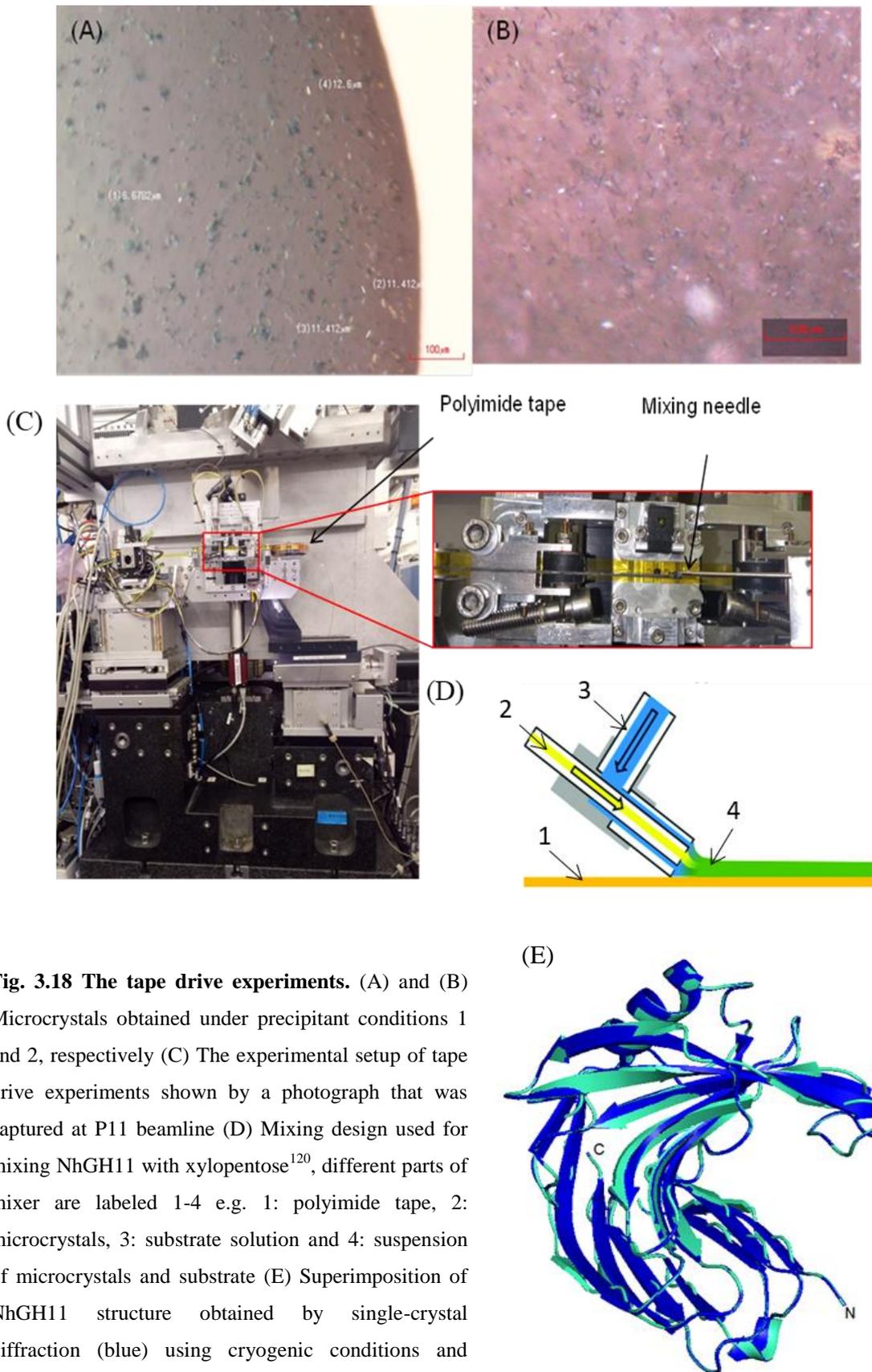


Fig. 3.18 The tape drive experiments. (A) and (B) Microcrystals obtained under precipitant conditions 1 and 2, respectively (C) The experimental setup of tape drive experiments shown by a photograph that was captured at P11 beamline (D) Mixing design used for mixing NhGH11 with xylopentose¹²⁰, different parts of mixer are labeled 1-4 e.g. 1: polyimide tape, 2: microcrystals, 3: substrate solution and 4: suspension of microcrystals and substrate (E) Superimposition of NhGH11 structure obtained by single-crystal diffraction (blue) using cryogenic conditions and structure obtained by data collection at room temperature (cyan).

Table 3.5 Room temperature data collection

X-ray source	P11 beamline PETRA III, DESY
Detector	Pilatus
Space group	C 2 (No. 5)
Cell dimensions	
a, b, c (Å)	80.55, 38.85, 53.57
α, β, γ (°)	90.00, 91.44, 90.00
Wavelength (Å)	0.97
Resolution range (Å)	16.87 -1.51 (1.56 -1.51)
Total number of reflections	25487 (1882)
Completeness (%)	93.09 (67.07)
Refinement	
Reflections used	24429 (1230)
Reflection used for R_{free}	1918 (105)
R_{work}	0.1963 (0.4)
R_{free}	0.2241 (0.4)
No. atoms	1665
Protein	1580
Ligand/ion	Not observed
Water	85
Average B-factor (Å²)	22.05
For macromolecule	21.52
For water	32.02
R.m.s deviations	
Bond lengths (Å)	0.015
Bond angles (°)	1.71
Ramachandran	
Favored (%)	95.7
Allowed (%)	4.3
Outliers (%)	0.00

*Data in highest resolution shells are shown in parenthesis.

3.7 Co-crystallization and soaking experiments

Crystallization experiments of NhGH11 with three saccharidic substrates: xylobiose (X2), xylotriose (X3) xylopentose (X5), and phenolic inhibitors: coumaric acid, ferulic acid, caffeic acid and cinnamic acid were performed. The co-crystallization experiments were performed under two different precipitant conditions by changing pH, concentrations of substrates/inhibitors and temperature (Fig. 3.19).

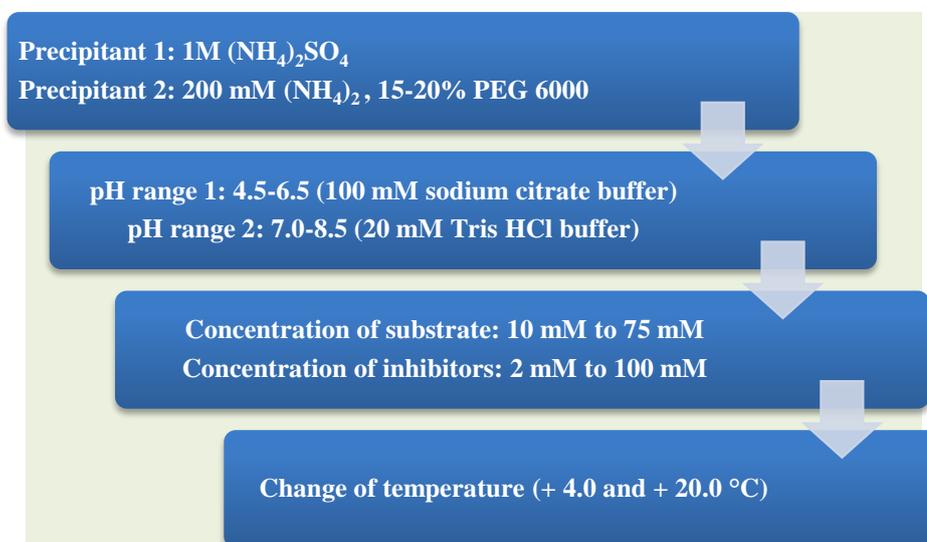


Fig. 3.19 Summary of physical and chemical variables for co-crystallization experiments of NhGH11.

The presence of substrates proved to be unfavorable for the crystallization process of NhGH11 since fewer crystals or no uniform shapes were obtained in the crystallization drops. Some successful conditions with precipitant 2 were obtained, in which NhGH11 crystals could grow in presence of saccharidic substrates (Fig. 3.20 A-E). However, these crystals did not show X-ray diffraction.

NhGH11 crystals can grow at maximum inhibitor concentrations of 7 mM, higher than 7mM inhibitor concentrations the crystals could not grow because of highly precipitated protein in the crystallization drops. Crystals grown with 7 mM inhibitors concentration are shown in Fig. 3.20 F-I, the diffraction data were collected to approx. 2.7 Å resolution, however, electron density was not observed for inhibitors.

Once unsuccessful results were obtained for tape drive and co-crystallization experiments, the soaking experiments were performed by applying saccharidic substrates (X2, X3) and phenolic inhibitors, the procedure described in section 2.13.3. Data collection and processing were performed by applying the standard procedures. Unfortunately, electron density was not observed for substrates or inhibitors.

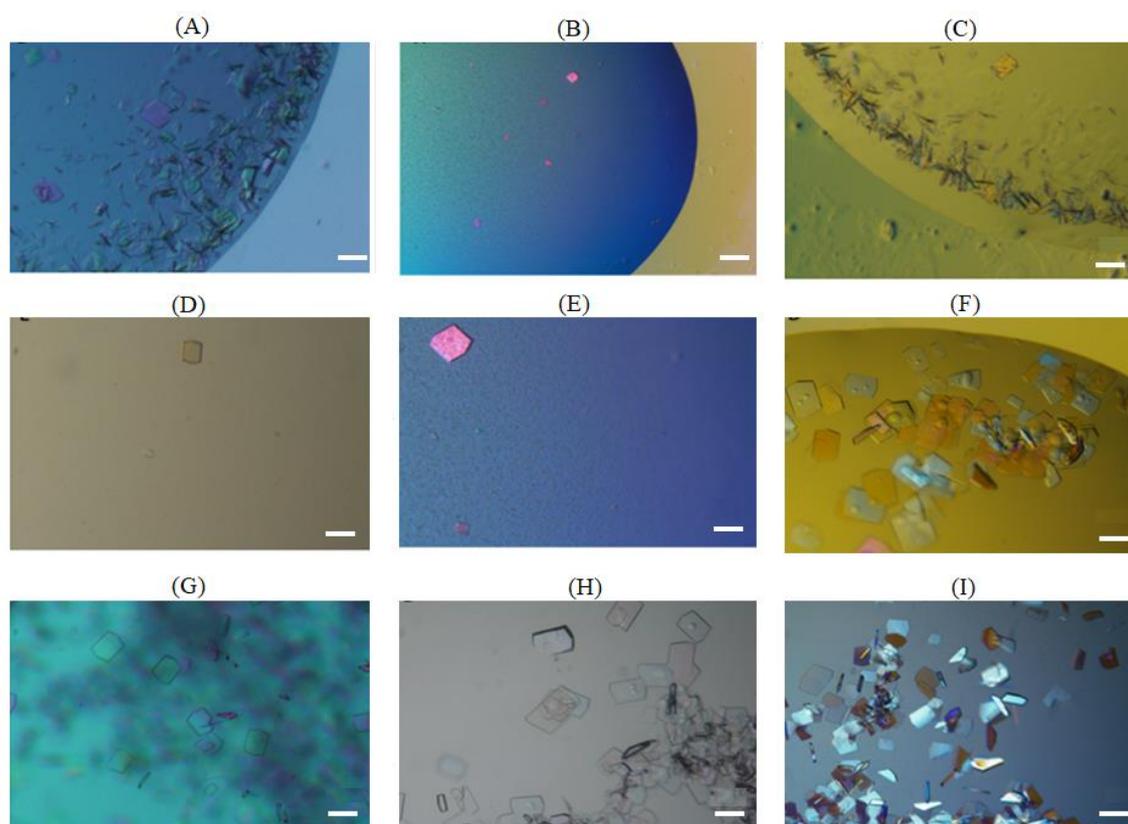


Fig. 3.20 Co-Crystallization with inhibitors and substrates. (A) Crystals were obtained in the presence of 75 mM xylobiose, 15% PEG 6000, 20mM Tris-HCl, pH 8.5. (B) Co-crystals obtained in the presence of 20 mM xylopentose, 15% PEG 6000, Sodium citrate 100 mM, pH 5.5 (C) Crystals obtained in the presence of 75 mM xylotriose with 15% PEG 6000, 20 mM Tris-HCl, pH 8.5. (D) Crystals obtained in the presence of 20 mM xylopentose with 20% PEG 6000, Sodium citrate 100 mM, pH 5.5. (E) Crystals were obtained in the presence of 20 mM xylopentose in presence of 20% PEG 6000, 20mM Tris-HCl pH 8.5. (F-I) Crystals were obtained in the presence of 7mM cinnamic acid, coumaric acid, ferulic acid and caffeic acid respectively with precipitant containing 1M $(\text{NH}_4)_2\text{SO}_4$, 100 mM sodium citrate pH 5.5. The scale bar corresponds to 200 μm .

The possible explanation of unsuccessful ligand binding experiments for NhGH11 can be the tight crystal packing as compared to ligand-bound xylanase from *T. reesei* (4HK8). The Matthews coefficient of the orthorhombic crystal form of 4HK8 is calculated to be 2.3 $\text{\AA}^3/\text{Da}$ that is higher than the monoclinic crystal form of NhGH11, which is 1.9 $\text{\AA}^3/\text{Da}$. Also, 4HK8 crystal contains higher solvent content (~47%) than NhGH11 crystal (~37%). Crystal packing showed that 4HK8 has broad intermolecular solvent channels compared to a tightly packed crystal of NhGH11. Crystal packing diagrams are shown in Fig. 3.21. In 4HK8 crystal, the active sites of molecules are facing towards broad solvent channels, while in the case of NhGH11 crystal the molecules are tightly packed in crystal allowing only a narrow active site cavity accessible to solvent as shown by a black arrow

in Fig. 3.21. In this arrangement of NhGH11 molecules, the substrate penetration could be hindered by the blockage of the active site channels of outside layers of molecules.

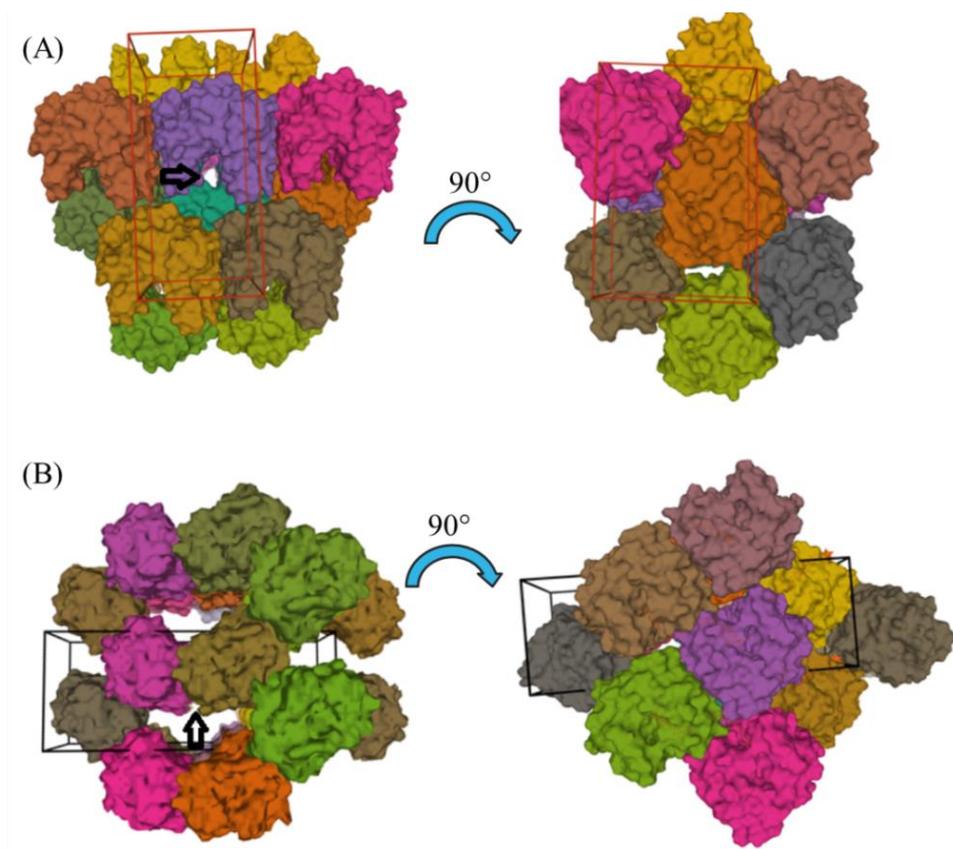


Fig. 3.21 Comparison of crystal packing. (A) Packing diagram for the monoclinic space group (No. 5) from NhGH11 crystal. (B) Packing diagram for the orthorhombic space group (No.18) from *T. reesei* xylanase crystal (PDB code 4HK8). Black arrows are indicating the position of the active site.

3.8 Biochemical and functional characterization of NhGH11

3.8.1 Optimum temperature

The curve of NhGH11 optimum temperature represents the effect of temperature, ranging from 30-80 °C, on the xylanase activity. NhGH11 showed activities scaled to $92.9 \pm 1.8\%$, $100.7 \pm 3.8\%$ and $84.8 \pm 2.7\%$ at 40 °C, 45 °C and 50 °C, respectively (Fig. 3.22). Relative activities were calculated by considering the activity to be 100% at optimum conditions. Optimum temperature and thermostability results showed that NhGH11 is very stable at 45 °C. These results are in good agreement with CD spectroscopy results which provide the evidence of 45 °C as the best working temperature and a mesophilic character of NhGH11.

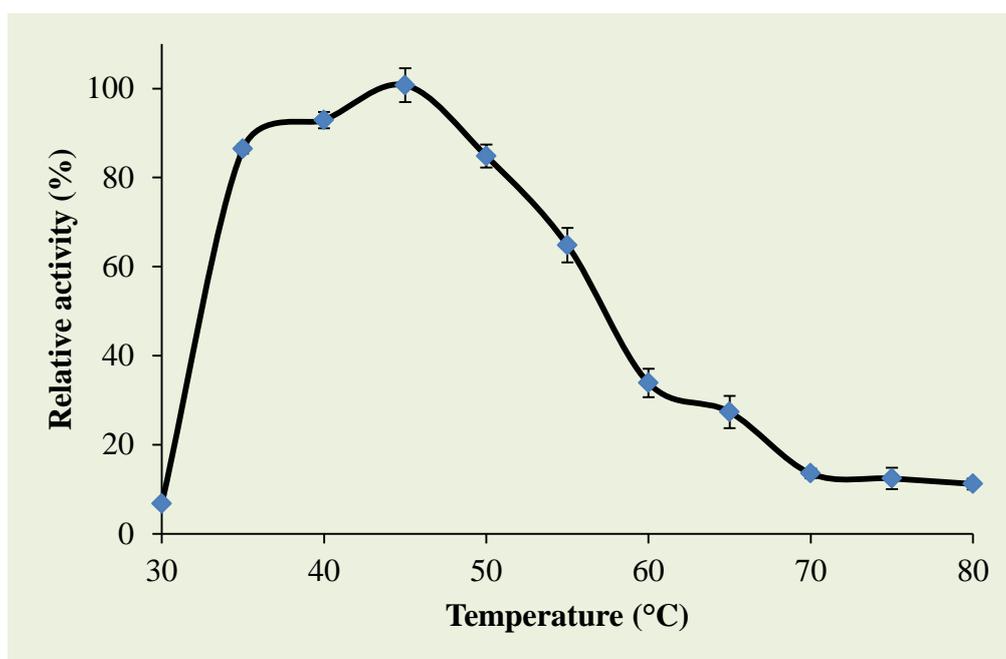


Fig. 3.22 Effect of temperature on enzyme activity of NhGH11. The optimum temperature for enzymatic activity was determined by measuring the activity at different temperatures. The data points represent the mean values of three replicates and error bars to indicate the corresponding standard deviation.

GH11 xylanases have a wide range of optimum temperatures²⁸. However, a calculated difference of energies between mesophilic and thermophilic enzymes was lower than 40.0 kJ mol^{-1} that can correspond to very few stabilizing interactions of 3D structure i.e., four hydrogen bonds considering 10 kJ mol^{-1} energy for each hydrogen bond. Consequently, the minor changes could provide major stabilizing interactions to enzymes¹⁹⁸. Optimum temperature decides a fine limit between stability, catalytic activity and inactivation of enzyme¹⁹⁹. When an active enzyme is subjected to progressive

temperature, in the first reversible step the non-covalent stabilizing interactions including hydrogen bonds and Van der Waals interactions break. At this stage, the enzyme becomes inactive due to the destabilization of non-covalent interactions and this inactivation is due to the partial unfolding of the enzyme. The effect of temperature on enzyme structure and function could be demonstrated by model²⁰⁰ shown below;



Further increase in temperature resulted in the exposure of protein core containing the hydrophobic residues to solvent-exposed that further lead to irreversible denaturation due to aggregation²⁰⁰.

3.8.2 Optimum pH

The effect of pH on the activity of NhGH11 was investigated to identify the optimum conditions for enzyme catalysis. NhGH11 showed 76.4±1.0%, 100.0±3.5% and 66.6±1.3% relative activities at pH 5.0, 6.0 and 7.0 respectively. Relative activities were calculated by considering the activity to be 100% at pH 6.0 and 45 °C as shown in Fig. 3.23.

The pH-dependent enzyme activity is primarily a set of pKa values due to ionic side chains of amino acids near to catalytic amino acid. GH11 xylanases have highly variable optimum pH values ranged from 2.0 to 9.0²⁸. Xylanase from *A. kawachii* is the most acidophilic enzyme with an optimum pH of 2.0 and even active at pH 1.0²⁰¹. Xylanase from *Bacillus* has a pH optimum of 9.0 and showed 60% activity between pH range from 5.0 to 9.5²⁰². The xylanases having isoelectric points less than 5.0 were considered as acidic and the remaining as alkaline²⁸.

Two xylanases from *H. jecorina* have two different pH optimum that was correlated with the type of residue at position 33 in the active site cleft²⁰³. The residue 33 could be Aspartate or Asparagine for acidophilic and alkaline xylanases. Because of the vicinity to the acid/base catalytic residue, the residue at position 33 could influence the ionization state of acid/base residue by creating a strong hydrogen bond and makes it easier to liberate a proton during catalysis²⁰⁴. The optimum pH could shift from acidic to alkaline conditions if this aspartate is mutated to an asparagine by site-directed mutagenesis²⁰⁵. Inspection of the 3D structure of NhGH11 showed that two asparagine residues are vicinal to catalytic glutamate 180, as shown in Fig. 3.24. Furthermore, the comparison of the 3D structure of NhGH11 and acidophilic *A. kawachii* xylanase (1BK1) showed that NhGH11

has asparagine while *A. kawachii* has aspartate vicinal to acid/base glutamate. Therefore, the structural evidence of alkaline pH of NhGH11 is correlated to the observed optimum pH.

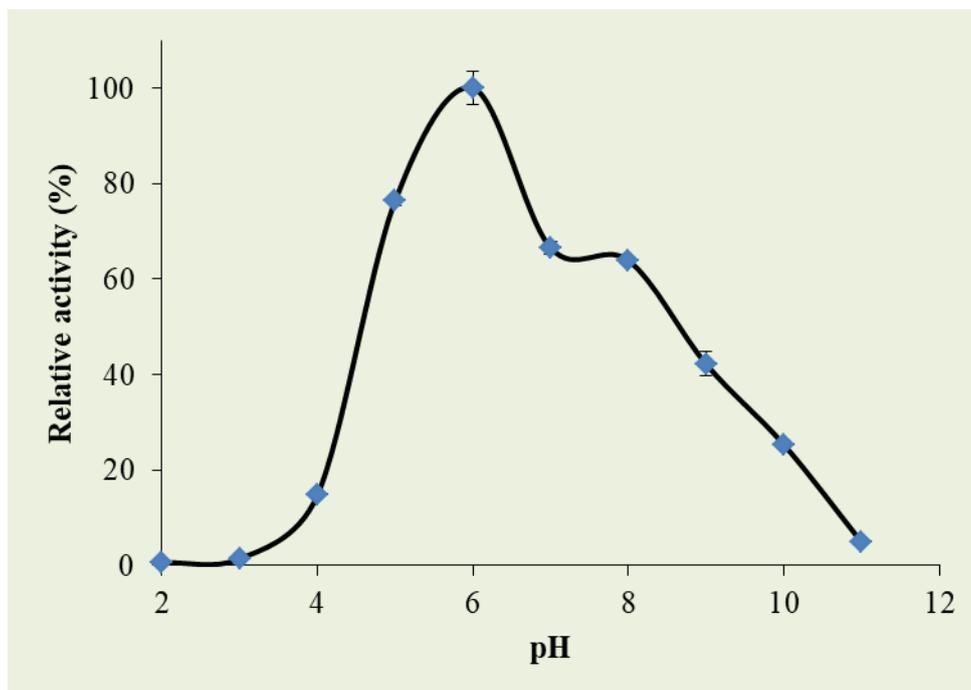


Fig. 3.23 Effect of pH on enzyme activity of NhGH11. The optimum pH for enzymatic catalysis was determined by using 100 mM McIlvaine buffers (pH 2-11). The data points represent the mean values of three replicates and error bars to indicate the corresponding standard deviation.

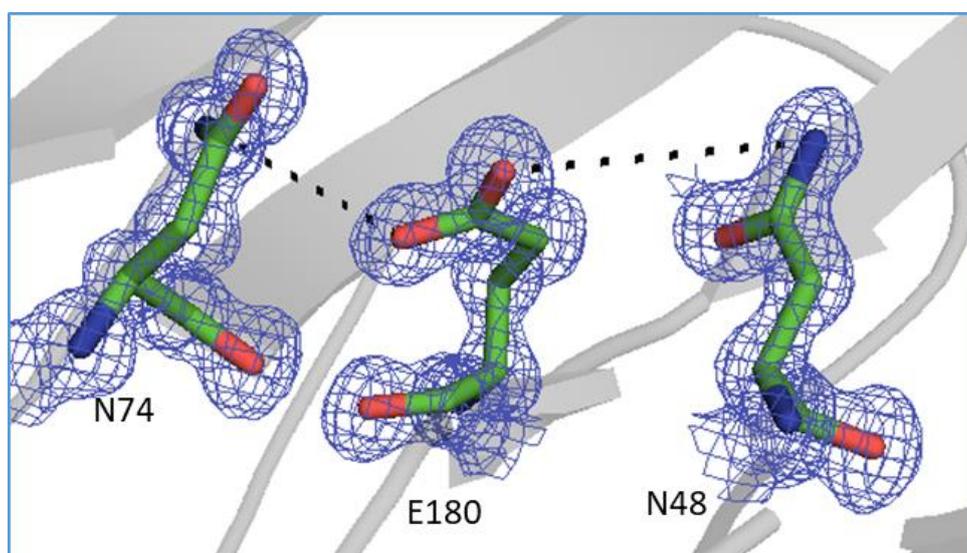


Fig. 3.24 Structural basis for optimum pH. Vicinal asparagine residues to catalytic E180 forming hydrogen bonds. 2Fo-Fc electron density maps are shown at 1 σ level.

Besides these, a reason for the low optimum pH of acidophilic GH11 xylanases could be the presence of acidic residues on the surface of acidophilic xylanases. The numerous acidic residues on the surface of xylanases are called Ser/Thr surfaces. At low pH, the acidic amino acids of acidophilic xylanases became protonated and neutral so these can easily accommodate the substrates without electrostatic repulsions. However, alkaline xylanases had less acidic residues on the surfaces which could correlate with stability against extreme pH conditions²⁰⁵.

Another important factor for optimum pH can be a higher number of ionic interactions for alkaline xylanases⁴⁰. Therefore, the analysis of ionic interactions was performed by applying a protein interaction calculator²⁰⁶. Ionic interactions were analyzed for acidophilic xylanases (optimum pH: 2.0) from *A. kawachii*²⁰¹, NhGH11 (pH: 6.0) and alkalophile xylanase (pH optimum: 8.5) from *Bacillus*²⁰⁷, the results of ionic interaction analysis showed that the acidophilic xylanase has only one ionic interaction, NhGH11 has three ionic interactions while highly alkalophilic xylanase has 17 ionic interactions.

Based on the comparison of optimum pH of acidophilic and alkalophilic xylanases, the ionic interactions, the nature of residues around the catalytic residues and the surface residues can be responsible for the pH stability, however, more investigations in terms of pH dependence on xylanase mechanism need to be performed²⁰⁸.

3.8.3 Effect of metal ions and chemical reagents

The effect of selected metal ions and chemical reagents was investigated for the activity of NhGH11 (Fig. 3.25). Na⁺ and K⁺ ions did not affect the enzyme activity of NhGH11. However, the presence of Ca²⁺, Co²⁺, Mg²⁺, Zn²⁺ and Mn²⁺ ions decreased the enzymatic activity scaled to the values e.g., 82.5±3.8%, 70.6±3.9%, 54.7±1.5%, 52.2±3.6% and 50.2±2.0%. While in presence of Fe²⁺ and Cu²⁺ ions, enzyme activity was more severely reduced to 25.6±4.3% and 15.6±2.6% respectively.

Pereira and collaborators (2017), reported the effect of metal ions on the activities of cellulases, hemicellulases and ligninases as shown in Fig. 3.26. They observed that some metal ions activated while others inhibited the cellulases, hemicellulases and ligninases. Furthermore, the number of charges and the ionic radius of metal ions also influenced the activities of glycosyl hydrolases²⁰⁹. The larger ionic radius influences less while the smaller ionic radius attracts the charged amino acids and damages the overall conformation of enzymes. Metals of group 2B e.g., zinc, cadmium and mercury showed a

high affinity for amino acids containing SH, COOH and NH₂ side chains²¹⁰.

The sensitivity of xylanases towards metal ions is not predictable among members of the GH11 family. For example, Mn²⁺ and Zn²⁺ enhanced the activity of *A. terreus* xylanase²¹¹ while Mn²⁺ and some other divalent ions slightly reduced the activity of *P. oxalicum* xylanase. However, for *P. oxalicum* xylanase, the metal ions activated the enzyme at lower concentrations but strongly deactivated it at higher concentrations²¹². The multiple xylanases from *P. oxalicum* were not substantially inhibited by the same metal ions. For example, the same concentration of Co²⁺ enhanced the activity of xylanase-A but strongly inhibited the xylanase-B²¹³. The xylanase from *T. halotolerans* showed that low concentration (1 mM) of Co²⁺ and Mn²⁺ activated the enzyme while Fe³⁺ and Pb²⁺ ions inhibited the activity at any concentration²¹⁴.

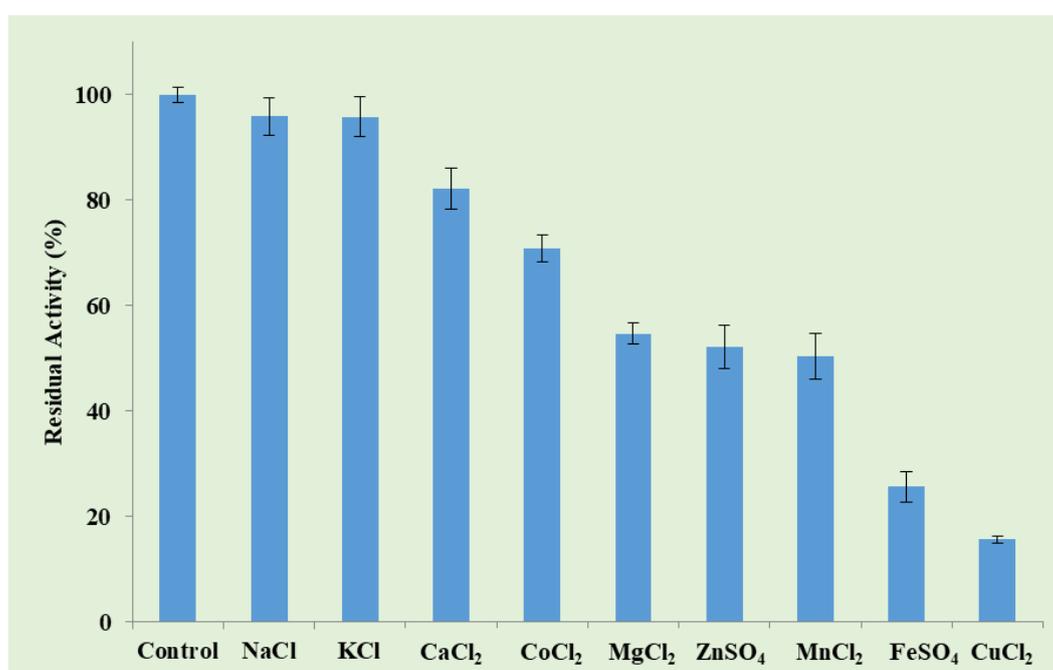


Fig. 3.25 Effect of metal ions on the activity of NhGH11. The data points represent the mean values of three replicates and error bars to indicate the corresponding standard deviation.

The effect of selected chemical reagents on the activity of NhGH11 showed that sodium azide, tween-80 and triton X-100 reduce the activity to 91.0±2.6%, 80.1±5.0% and 70.2±4.9% respectively (Fig. 3.27). NhGH11 showed a significant decrease in activity in the presence of EDTA and SDS. The decreased enzyme activity in the presence of EDTA, as a metal chelating agent, indicated that at low concentrations some divalent metal ions are supporting the activity of NhGH11. The effect of ionic detergents on arabinofuranosidase showed that at low concentrations (1-2 mM) some ionic detergents

including SDS enhanced the enzyme activity while at higher concentrations (>20 mM) strong inhibition was observed²¹⁵. Later on, SDS was found to interfere with hydrophobic regions of enzymes that resulted in denaturation at higher concentrations²¹⁶.

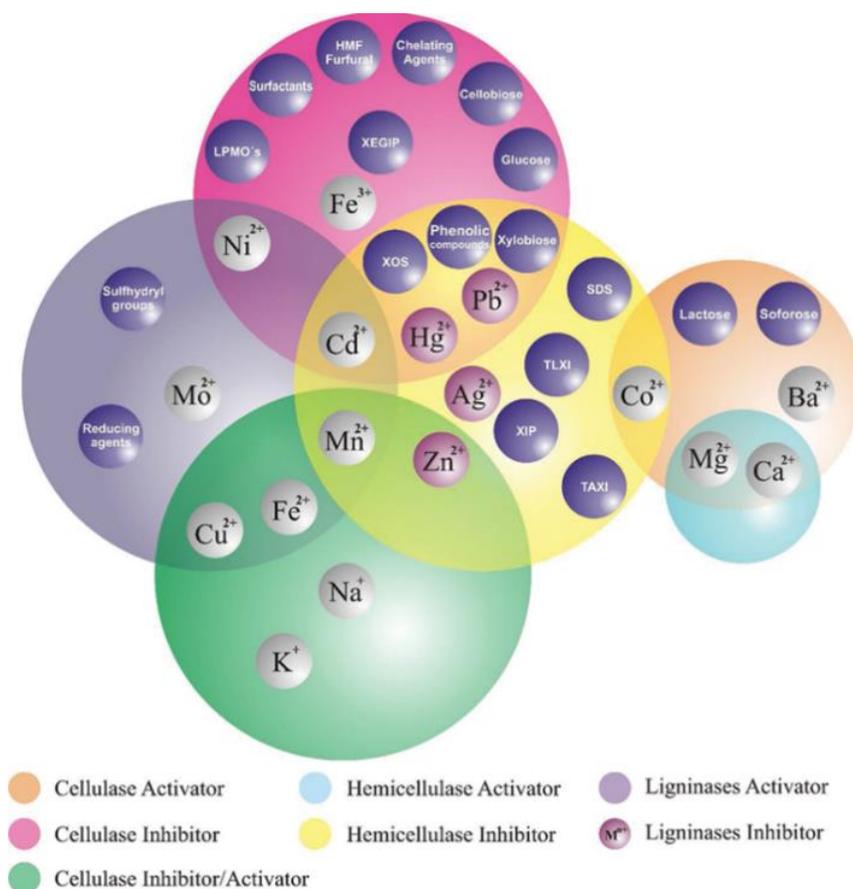


Fig. 3.26 Activators and inhibitors of GH enzymes. Hemicellulase inhibitors (yellow) and activators (cyan) also affect the activities of cellulases and ligninases²¹⁰.

The effect of surfactants on the activity of *P. oxalicum*²¹² and NhGH11 showed that these enzymes are more sensitive towards SDS followed by EDTA and β -mercaptoethanol as compared to tween-80 and triton X-100. The effect of chemical reagents on the activity of multiple xylanases from *P. oxalicum* showed that multiple xylanases from the same fungi showed different effects for chemical reagents. For example, DTT, triton and tween-20 did not influence much the enzymatic activity for all xylanases from *P. oxalicum*. EDTA and β -mercaptoethanol strongly inhibited the enzymatic activity of xyn11A from *P. oxalicum* but the xyn10A, xyn10B and xyn11B were only slightly inhibited²¹³.

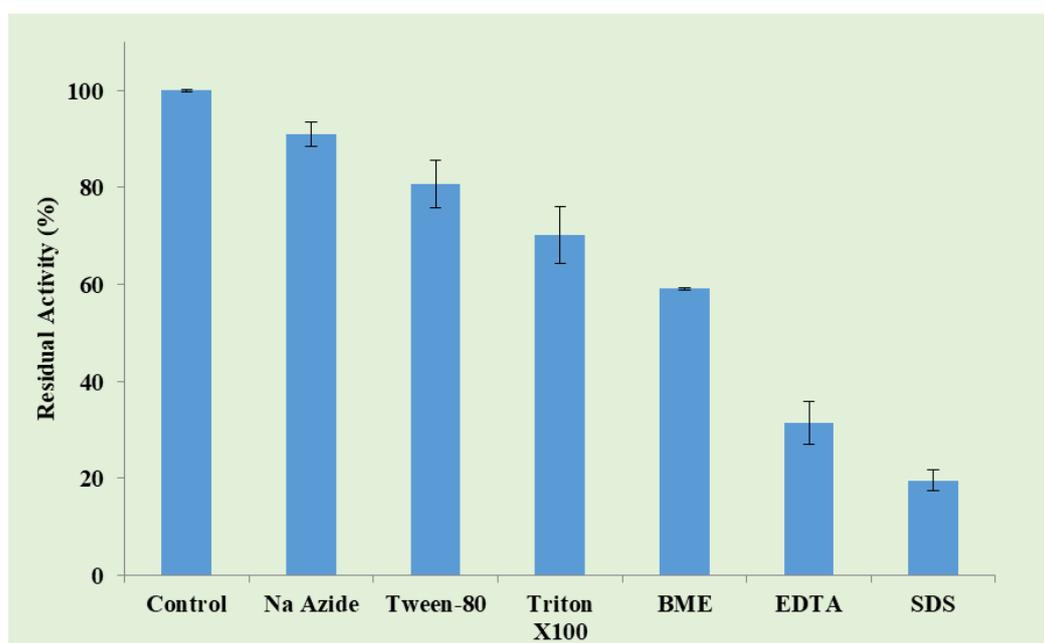


Fig. 3.27 Effect of chemical reagents on the activity of NhGH11. The data points represent the mean values of three replicates and error bars to indicate the corresponding standard deviation.

3.8.4 Substrate specificity and kinetic parameters

Kinetic parameters were analyzed considering the pre-determined optimum conditions of NhGH11 by applying different substrates. The catalytic activity of NhGH11 was tested towards xylan beechwood, xylan polysaccharide, wheat arabinoxylan, azo-xyloglucan and mannan polysaccharide. All tested substrates are structurally different e.g. xylan beechwood is β -1,4 linked xylan with 8.7% of 4-O-methyl glucuronic acid and 5.7% other sugars (Sigma, USA). Xylan polysaccharide is a cell wall polysaccharide from *Palmaria palmate* that is edible seaweed. It is β -1,3 and β -1,4-linked xylose polymer (Elicityl, France). Wheat arabinoxylan is a pentosan fraction from wheat flour. The main chain is β -1,4 linked xylan with the branches of 33.5% arabinose, 62% xylose, 1.5% glucose, 2.0% mannose and 1.0% galactose. Azo-xyloglucan from tamarind contained β -1,4 linked glucose with branches of galactose, xylose, arabinose and other sugars. (Megazyme, Australia). While mannan polysaccharide contained β -1,4-linked mannose with 2% galactose and other sugars (Sigma, USA). The chemical structures are shown only for substrates on which NhGH11 showed catalytic activities (Fig. 3.28).

Activity graphs were plotted based on Michaelis-Menten kinetics, used to calculate the K_m and V_{max} for each substrate (Fig. 3.29). NhGH11 catalyzed the hydrolysis of xylan beechwood, xylan polysaccharide and wheat arabinoxylan. The specificity constant, k_{cat}/K_m ($\text{mg}^{-1} \text{ml min}^{-1}$), applying xylan beechwood as substrate was higher than for wheat

arabinoxylan and xylan polysaccharides, while the turnover number, k_{cat} (min^{-1}), for wheat arabinoxylan was higher than for xylan beechwood and xylan polysaccharides respectively (Table 3.6). NhGH11 was found to be active towards xylan substrates e.g., xylan beechwood, xylan polysaccharide and wheat arabinoxylan but not towards azo-xyloglucan and mannan.

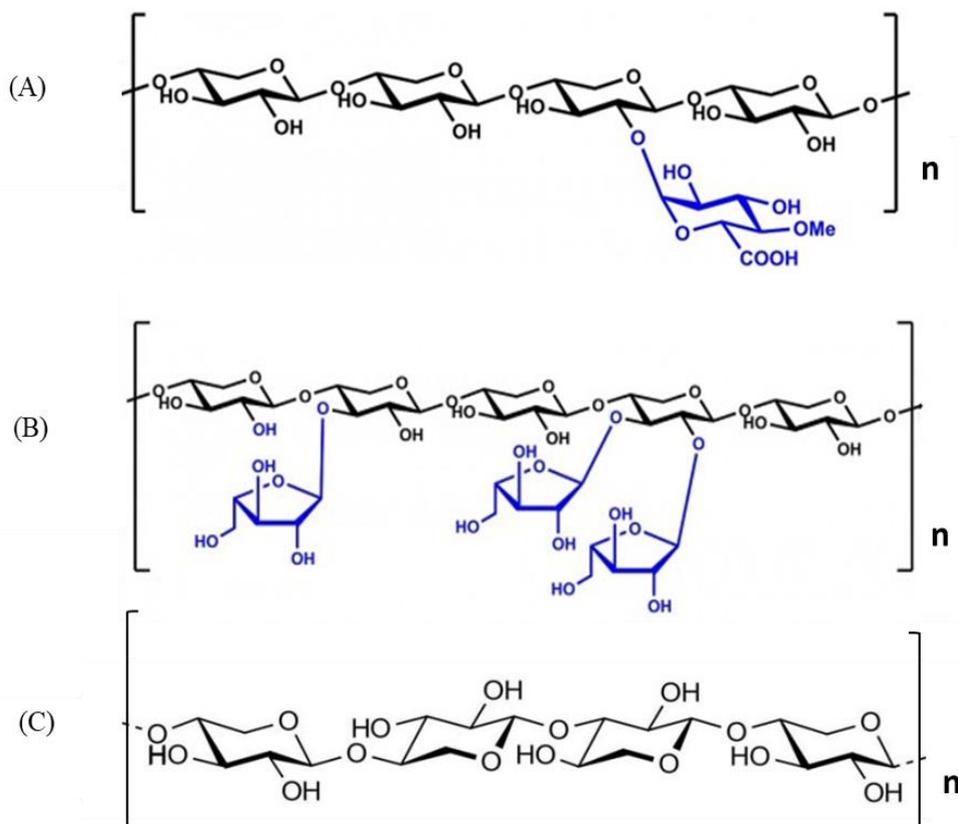


Fig. 3.28 Chemical structures of Xylan substrates. (A) Xylan beechwood with branches of 4-O-methyl glucuronoxylan, (B) xylan polysaccharide from *Palmaria palmate* (C) wheat arabinoxylan.

Table 3.6 Comparison of kinetic values of NhGH11 with various substrates

Xylan substrates	K_m (mg/ml)	V_{max} ($\mu\text{M}/\text{min}$)	k_{cat} (min^{-1})	k_{cat} / K_m ($\text{mg}^{-1} \text{ml min}^{-1}$)
Xylan beechwood	8.1 \pm 2.1	14.1 \pm 2.6	1730.6 \pm 318.1	214.9 \pm 68.2
Xylan polysaccharide	10.1 \pm 2.3	13.4 \pm 2.0	1648.2 \pm 249.2	163.7 \pm 44.7
Wheat arabinoxylan	13.0 \pm 4.1	19.7 \pm 4.2	2410.8 \pm 517.5	200.9 \pm 81.1

Mostly, GH11 xylanases are known for the higher specific activity for xylyans in comparison to GH10 xylanases, which can hydrolyze a broad range of different xylan and cellulose substrates²¹⁷. GH11 xylanases showed specific endo- β -1,4-xylanase activities but GH10 showed promiscuous activities including endo- β -1,4-glucanase, endo- β -1,4-xylanase and endo- β -1,3-xylanase activities²¹⁸. Multiple xylanases from *P. oxalicum* were studied for comparative analysis of xylanase activities. All multiple xylanases from GH11 showed activities towards xylyans from oat spelt, birchwood and beechwood while GH10 xylanase showed activity towards commercial celluloses because of promiscuous activities of GH10 xylanases²¹³. Xylanases from the GH11 family have a higher affinity towards xylan beechwood. For example, NhGH11 and GH11 xylanase from *A. niger* showed the lowest K_m for xylan beechwood, indicating a higher affinity²¹⁹. Also, GH11 xylanases from *Bispora* showed the highest specific activity for xylan beechwood compared to xylyans from oat spelt and birchwood²¹². While, xylanases from *A. sydowii* also showed activity for xylan substrates e.g., xylan birchwood and oat spelt but not towards other substrates like pectin and starch²²⁰.

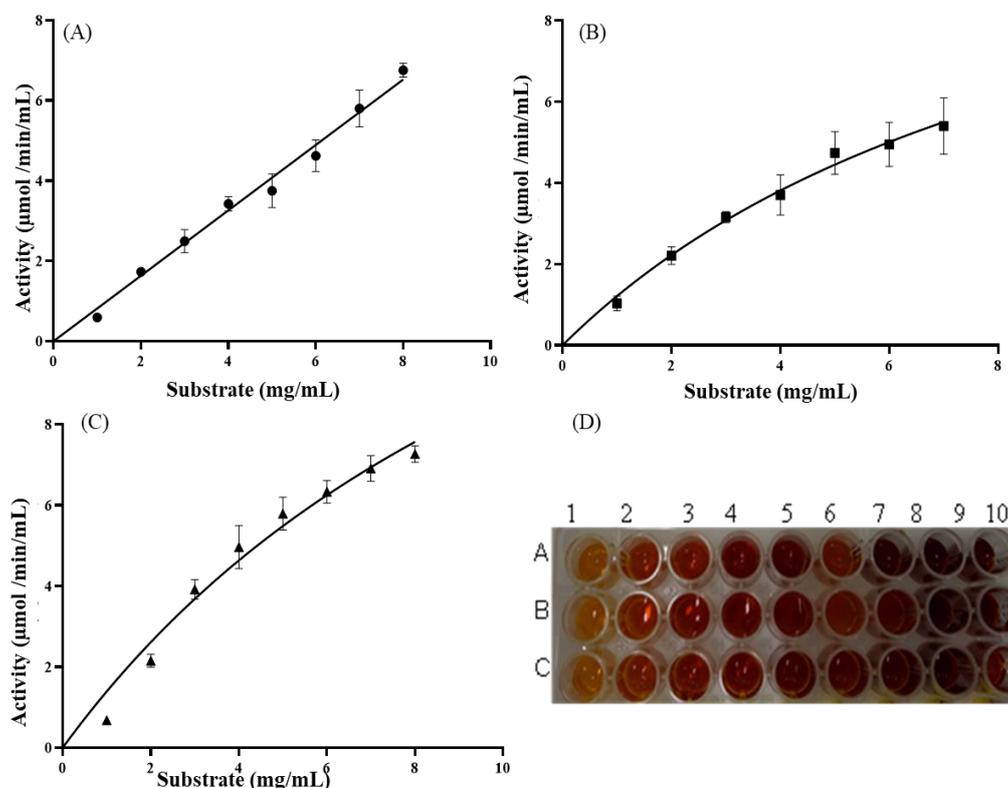


Fig. 3.29 Substrate specificity and kinetic parameters. The activity of NhGH11 was observed for substrates; (A) Xylan beechwood (B) Xylan polysaccharide from *Palmaria palmate* (C) Wheat arabinoxylan. (D) Spectrophotometric plate indicating the triplicate measurement of NhGH11 activity using xylan beechwood.

3.8.5 Identification of products

The xylooligosaccharides (XOSs) are oligomers of xylose sugar that are linked through β -1,4 xylosidic linkages forming xylobiose, xylotriose, xylotetrose and xylopentose with 2, 3, 4 and 5 monomers respectively. The products of the catalytic breakdown of xylopentose by NhGH11 were analyzed by MALDI-TOF-MS (matrix-assisted laser desorption ionization-time of flight mass spectrometry).

The molecular weights of XOSs were calculated by the addition of the molecular weight of sodium ions (Fig. 3.30). The spectra peaks between 100-800 Da were considered for the identification of products. Most prominent peaks correlated to xylobiose (304 Da), xylotriose (436 Da), xylotetrose (568 Da) and xylopentose (700 Da). The results showed that NhGH11 cleaved the longer xylan polymer into smaller XOSs like Xylose (X), Xylobiose (X2), Xylotriose (X3) and Xylotetrose (X4). The corresponding peak of Xylose was not identified in the spectra. The presence of X4 provided evidence of xylose. The proposed cleavage mechanism of xylopentose is shown in Fig. 3.31.

GH families of xylanase have different substrate-binding sites and specificities resulted in different ratios of final products²²¹. For example, GH8 xylanase from *T. turnerae*²²² cannot cleave the XOS product less than X4, while GH11 xylanase from *P. oxalicum*²¹² can hydrolyze smaller XOS as well. Compared to other xylanases, NhGH11 can produce industrially important smaller XOSs. XOSs can be used as food ingredients and prebiotics for the maintenance of intestinal function²²³. Furthermore, XOS were categorized as nutraceuticals because of their potential application in the prevention of atherosclerosis, they also show anti-inflammatory, anti-hyperlipidemia and anti-cancer activities²²⁴.

XOSs with glucuronate arabinose branches were reported to inhibit the growth of sarcoma and tumors²²⁵. Moreover, XOSs can also play an important role in inhibiting breast and colon cancers²²⁶. XOSs from almond shells were studied for their immunomodulation effects in rats also O-acetylated and de-acetylated XOSs showed the mitogenic effect to enhance the proliferation of thymocytes²²⁷. XOSs were found to be very effective to inhibit the growth of *S. enteritidis*²²⁸.

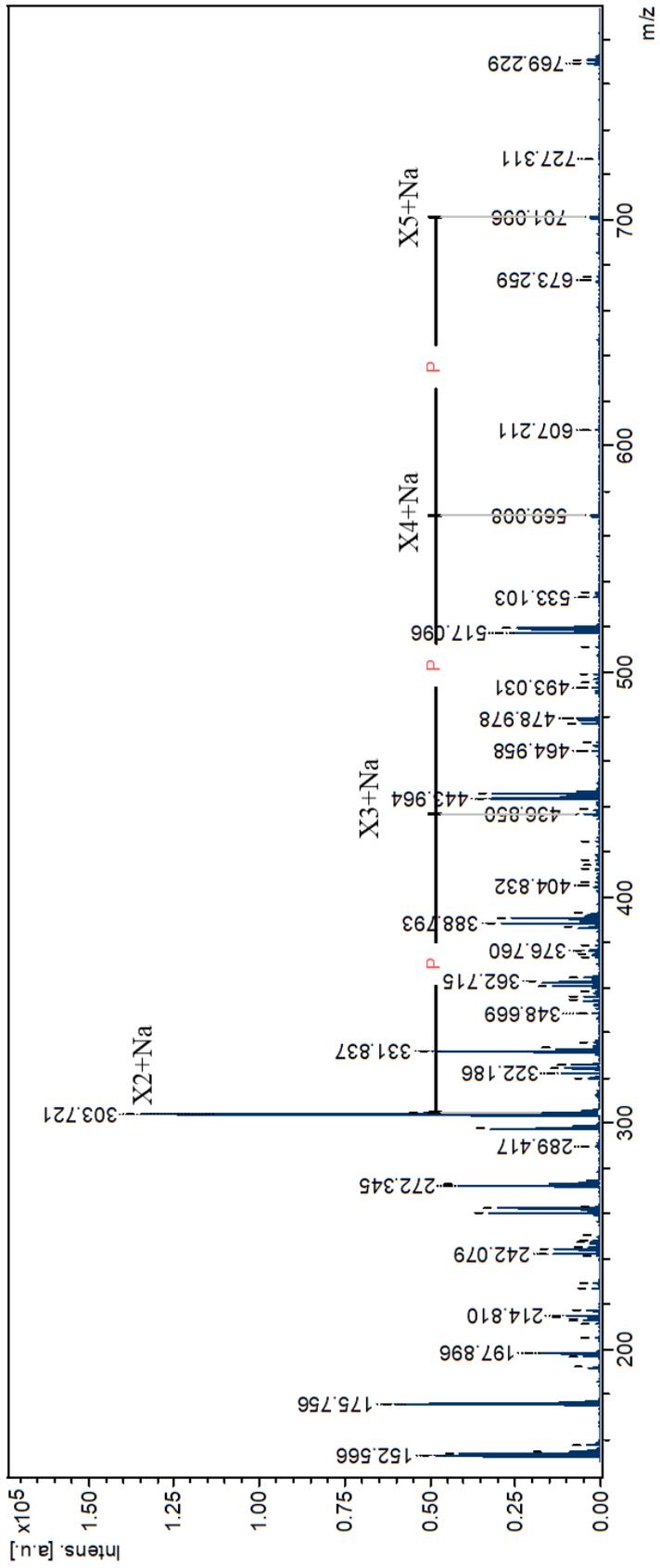


Fig. 3.30 Identification of catalysis products. MALDI-TOF MS spectrum of products (X2, X3, X4) formed by cleavage of substrate (X5) measured by using DHB matrix, the above green spectra showed the blank measurement of matrix and below blue spectra represents the xylooligosaccharides adducts with sodium ions.

XOSs showed a potent antioxidant activity and can be used to protect against lipid peroxidation²²⁹. XOSs with 2-10 units of xylose sugar cannot be digested like food whereas they showed different health benefits and physiological functions. Besides these biological functions, XOSs have also phytopharmaceutical applications^{64,224}. Owing to the health benefits of XOSs, they can be used as food additives in chocolate, bread and biscuits with a safe range of concentration between 7-10 percent⁶⁴. The permissible concentration for the human body is 0.12 g/kg of body weight²³⁰. Therefore, XOSs products of NhGH11 can be applied for similar applications.

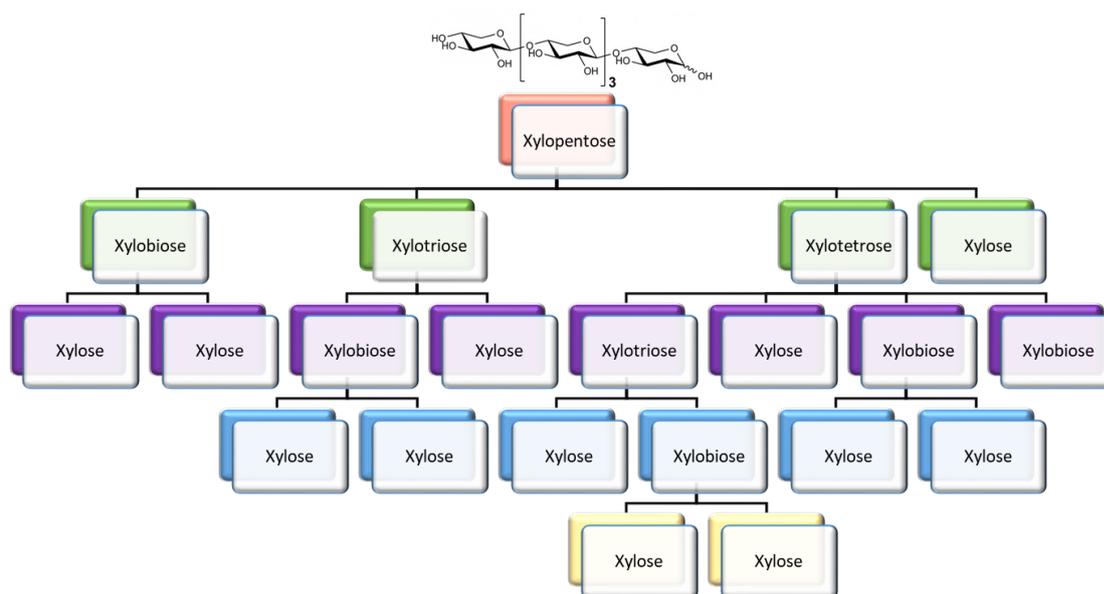


Fig. 3.31 The proposed mechanism of xylopentose hydrolysis. NhGH11 catalyzed the hydrolysis of longer xylan polymers into smaller xylooligosaccharides. The proposed mechanism includes the cleavage of xylopentose into xylo-tetrose and xylose, xylo-tetrose into xylo-triose and xylose, xylo-triose into xylo-biose and xylose.

3.9 Inhibition studies

3.9.1 Residual activity

The inhibitory action of lignin-derived phenolic compounds i.e. caffeic acid, cinnamic acid, coumaric acid and ferulic acid were comparatively investigated towards xylanase activity of NhGH11. The detrimental effect of lignin-derived phenolic compounds on the activity of NhGH11 is shown in Fig. 3.32. None of these phenolic compounds showed considerable inhibition at concentrations up to 1 mM. However, above 1.5 mM, a concentration-dependent decrease of residual activity was observed for all phenolic

compounds reaching 20-30% of residual activity at 4.5 mM.

These phenolic compounds also inhibited the GH11 xylanase from *T. xylanilyticus*²³¹. Additionally, the lignin-derived phenolic compounds were also found as growth inhibitors of fungi that secrete several GH enzymes for biofuel industries. As an example, the growth rate of *A. japonicus* was significantly inhibited in the presence of various concentrations of phenolic compounds as vanillin, cuaiacol and caffeic acid the growth. Also, the results of in vitro assays showed that these phenolic compounds inhibit the arabinases and xylanases from *A. japonicas*²³². Xylanase from *E. nidulans* was observed to be less sensitive towards phenolic inhibitors i.e. vanillin, ferulic acid, coumaric acid and 4-hydroxybenzoic acid, however, cinnamic acid and tannic acid inhibited it strongly. Inhibition of xylanases by phenolic compounds draws attention towards the investigation of their effects on particular xylanases to be applied in the industrial process²³³. The identification of xylanase inhibitors and the mechanism of inhibition can provide useful information to moderate the enzymatic hydrolysis of lignocellulosic biomass for biofuel production¹⁴⁵.

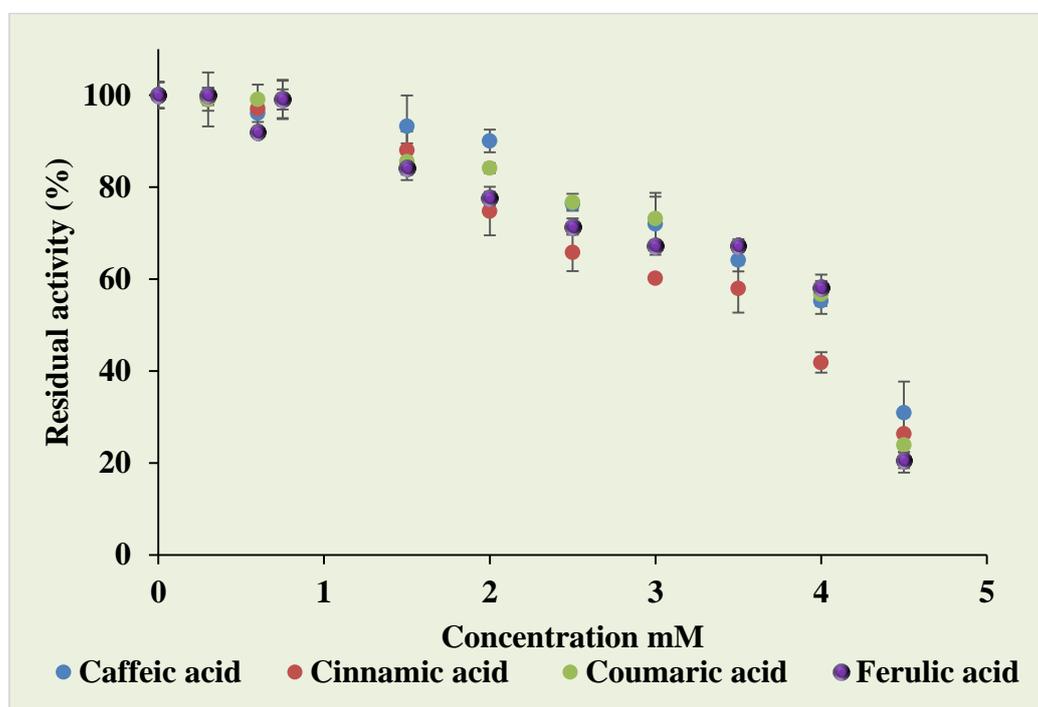


Fig. 3.32 Residual activity of NhGH11 in presence of lignin-derived inhibitors. The residual activity was calculated in the presence of lignin-derived inhibitors as caffeic acid, cinnamic acid, coumaric acid and ferulic acid. The activity of the purified enzyme was taken as 100%. The data represent a mean value of three replicates and bars indicate the standard deviation of the three replicates.

3.9.2 Thermofluor assay

The thermofluor assay is used to calculate the thermal unfolding of proteins in presence of ligands and monitors the effects of ligands on temperature-dependent protein unfolding²³⁴. In this assay, SYPRO orange dye was used which binds to hydrophobic regions during the thermal unfolding of proteins. The dye gives fluorescence after interacting with hydrophobic regions of proteins and as a result, the obtained sharp sigmoidal melting curve allows the calculation of the melting temperature (T_m) that corresponds to the 50% unfolded fraction of protein, as shown in Fig. 3.33.

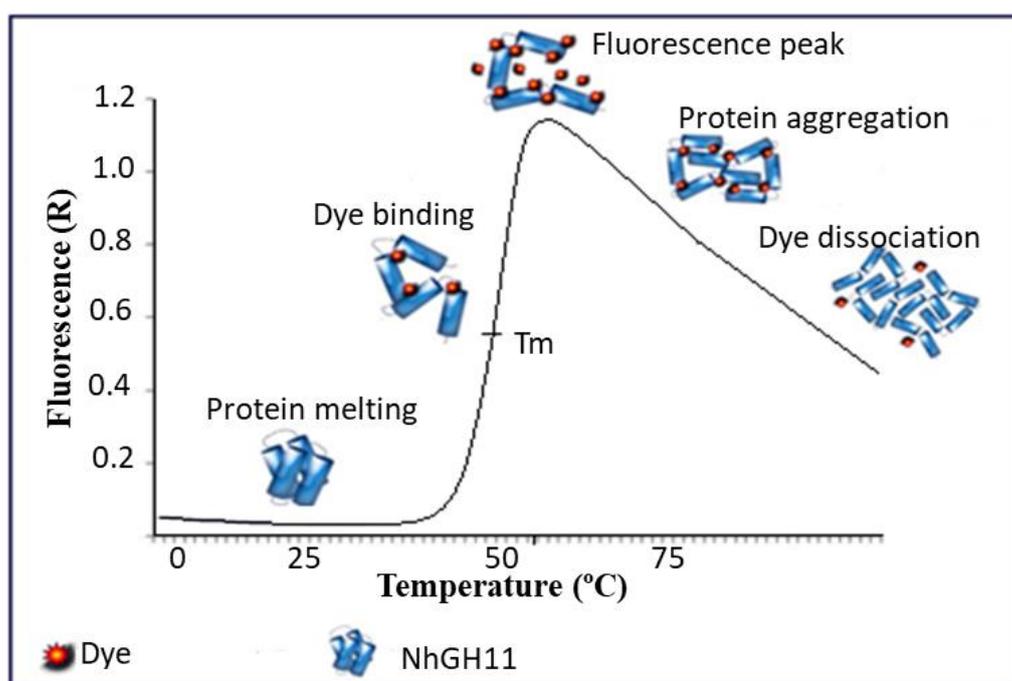


Fig. 3.33 Thermofluor assay. Temperature unfolds the protein and the exposed hydrophobic regions bind to the dye, resulting in a fluorescence peak. Figure source: <https://www.jenabioscience.com/crystallography-cryo-em/screening/thermofluor-screens>.

Reinhard and collaborators (2013), reported that thermofluor can be used to monitor the stability and homogeneity of proteins²³⁵. Thereby, the stability of proteins can be observed by the positive value of melting temperatures ($+\Delta T_m$), which can be evidence of increased structural stability and less conformational flexibility. On the other hand, the negative value ($-\Delta T_m$) can indicate a disordered conformation or a misfolded protein structure²³⁶. NhGH11 thermofluor curves are shown in Fig. 3.34.

The melting temperature ($T_m = 50.7\text{ °C}$) of NhGH11 without phenolic compounds, which was assumed as a control, was used to calculate the change in melting temperature (ΔT_m)

in presence of the phenolic compounds. The most prominent destabilization was observed after adding 600 μM of phenolic compounds, mainly in the presence of cinnamic acid which promoted a faster initial protein unfolding at a temperature of approx. 34 $^{\circ}\text{C}$, while 600 μM of caffeic, ferulic and coumaric acids showed initial unfolding at approx. 41-42 $^{\circ}\text{C}$.

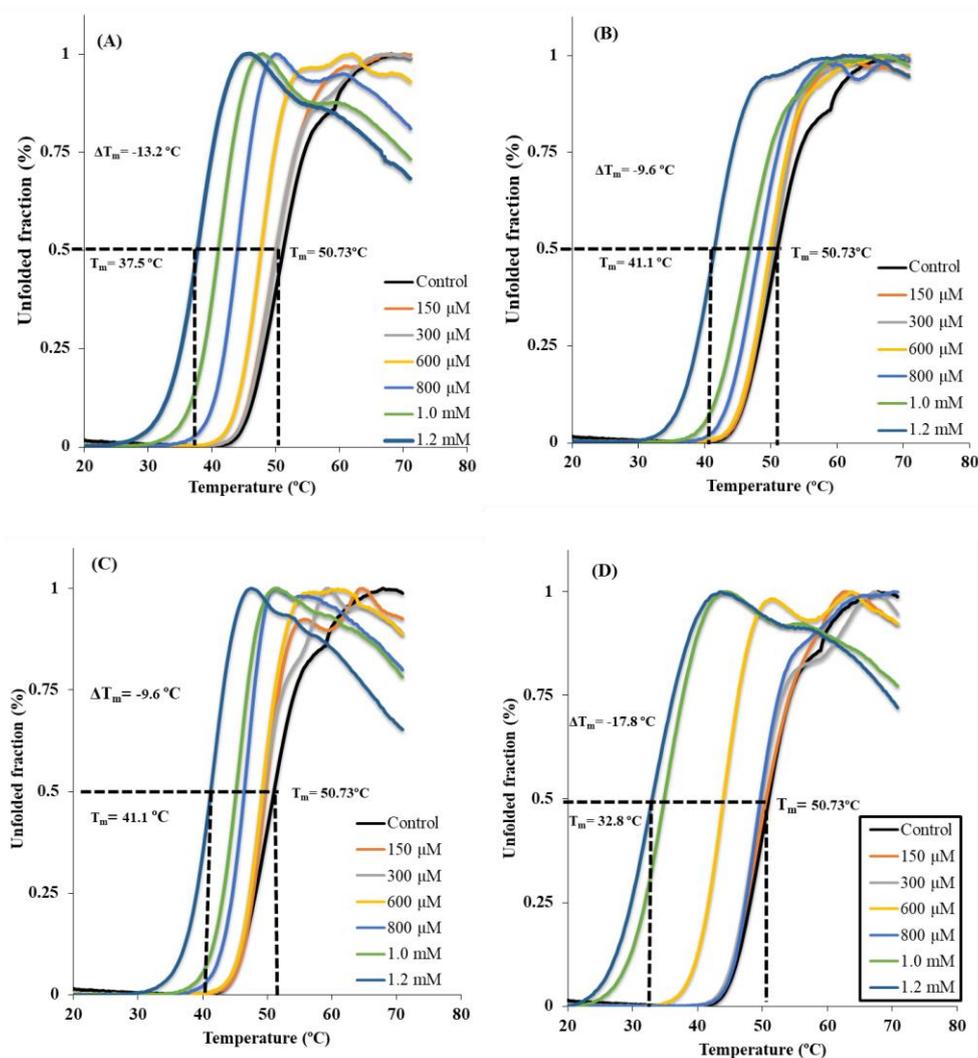


Fig. 3.34 Effect of phenolic compounds on the stability of NhGH11. Measurements were done on a real-time PCR machine using a temperature gradient of 1 $^{\circ}\text{C}/\text{min}$. (A) coumaric acid (B) ferulic acid (C) caffeic acid (D) cinnamic acid.

NhGH11 destabilization is a concentration-dependent effect (Fig 3.35) observed for all phenolic compounds and revealed the stronger effect in presence of cinnamic acid ($\Delta T_m = -17.8$ $^{\circ}\text{C}$), coumaric acid ($\Delta T_m = -13.2$ $^{\circ}\text{C}$), followed by ferulic and caffeic acid (both $\Delta T_m = -9.6$ $^{\circ}\text{C}$) and the two last showed less structural destabilization of NhGH11.

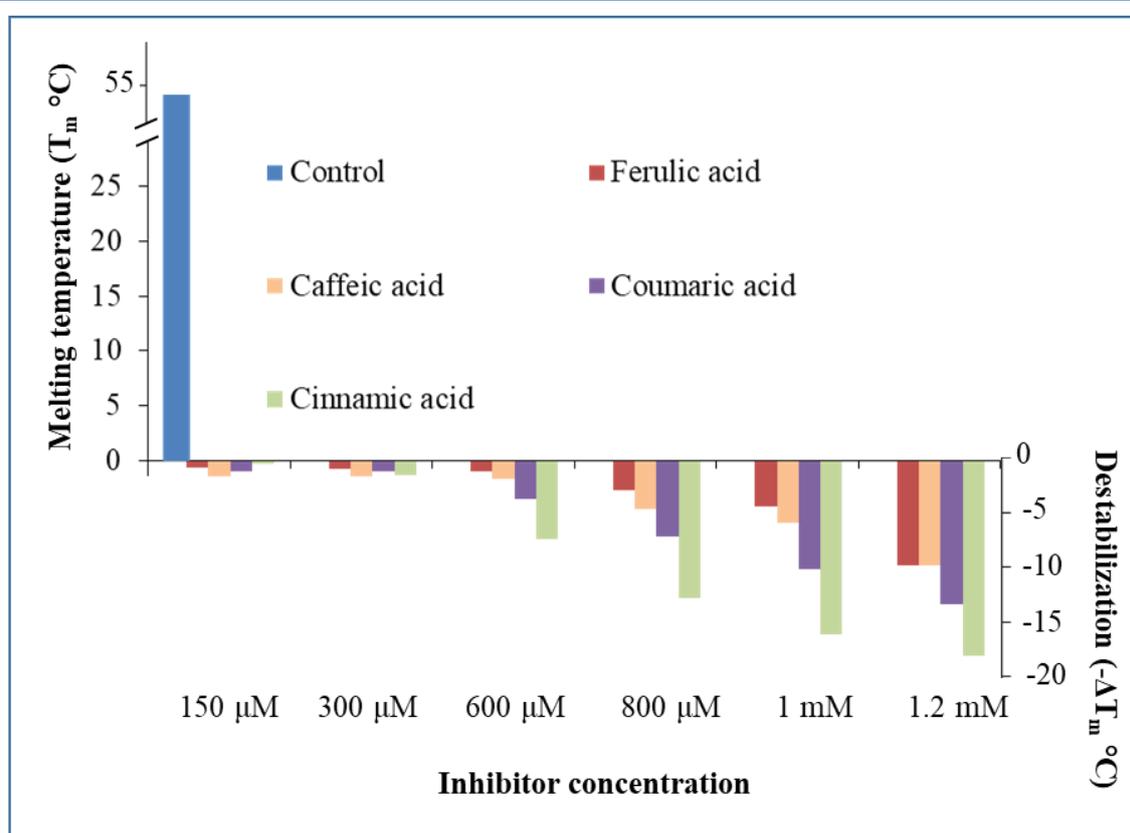


Fig. 3.35 Destabilization effect in presence of phenolic compounds. The histogram indicates the melting temperature of NhGH11 as control and change in melting temperature (ΔT_m) with increasing concentrations of phenolic compounds from 150 up to 1.2 mM.

Protein surfaces tend to bind organic molecules that can promiscuously cause inhibition by partial unfolding of the protein structure²³⁷. The predictions of binding pockets of the NhGH11 structure (Fig. 3.36 A) and calculations of possible interacting residues (Table 3.7) were obtained by applying DoGSiteScorer, ProteinsPlus web service (<https://proteins.plus>)²³⁸.

The possible interactions of phenolic compounds with NhGH11 can be explained by a comparison of hydrogen bond interaction made by coumaric acid in the crystal structure (PDB code: 4EYO) of ABC transporter from *R. palustris* (Fig. 3.36 B). The comparison with the protein-coumaric acid interactions for PBD structure 4EYO is indicating that the hydroxyl group of phenol can form hydrogen bonds with nitrogen atoms of glutamine and histidine. While the carboxyl group can form a hydrogen bond with positively charged residues and with hydroxyl groups of vicinal side chains. Additionally, the benzene ring of the phenolic compound can form aromatic-aromatic (π - π) interactions and C-H- π interactions with aromatic residues. Similar interaction can be observed for NhGH11 binding pockets with charged residues and H-bond donors and acceptors (Table 3.7).

Table 3.7 Description of predicted binding sites of NhGH11

Binding site	Red	Green	Blue	Yellow	Purple
Size and shape descriptors					
volume [\AA^3]	118.27	135.23	107.52	284.67	138.11
surface [\AA^2]	160.10	217.28	242.77	380.42	145.87
depth [\AA]	10.01	10.42	9.45	13.91	8.85
Functional group descriptors					
No. of H- bond donors	5	5	5	10	5
No. of H- bond acceptors	18	16	14	20	24
No. of hydrophobic interactions	9	7	14	11	5
Amino acid composition					
Apolar amino acid ratio	0.18	0.17	0.25	0.43	0.28
Polar amino acid ratio	0.82	0.75	0.33	0.48	0.61
Positive amino acid ratio	0.00	0.00	0.25	0.05	0.11
Negative amino acid ratio	0.00	0.08	0.17	0.05	0.00

Additionally, the residues of the red binding pocket can form more hydrogen bonds due to its polar amino acid composition followed by green, purple, yellow and blue pockets. These interactions may explain the possible reason for the instability of NhGH11 in presence of phenolic compounds including; the formation of hydrogen bonds between protein-phenolic and consequently the disruption of the protein-protein hydrogen bonds in the binding locality. The mechanism of protein destabilization in presence of ligands is not well understood due to lack of proper methodology. However, McGovern and collaborators (2002) suggested, that the query of destabilization of proteins in presence of ligands can be addressed by a combination of biophysical studies, fluorescence spectroscopy and kinetic assays to correlate the mechanism of inhibition²⁴⁰. Therefore, tryptophan fluorescence spectrophotometry was applied to find the effect of phenolic compounds on tryptophan fluorescence spectra of NhGH11 and the calculation of binding constants.

3.9.3 Tryptophan fluorescence spectrophotometry

The tryptophan fluorescence spectrophotometry was used to measure the intrinsic fluorescence of tryptophan residues and also the quenching of fluorescence in the presence of ligands²⁴¹. Tryptophan fluorescence spectra of proteins can provide evidence of changes in the tryptophan environment by ligand binding²⁴². Therefore, the effect of phenolic compounds on tryptophan fluorescence spectra of NhGH11 was investigated. The three-dimensional structure of NhGH11 contains six tryptophan residues at positions 21, 42, 55, 82, 141 and 162. Tryptophan residues at positions 21, 82 and 141 are partially solvent-exposed in the active site cleft while residue 55 is solvent exposed near to surface. However, W162 and W42 are buried in the hydrophobic environment (Fig. 3.37 A).

Buried and solvent-exposed tryptophan residues contribute unequally to the tryptophan emission spectra. For example, the spectra of apoazurin, ribonuclease, staphylococcal nuclease and glucagon are shown in Fig. 3.37 B²⁴³. In apoazurin and ribonuclease, the tryptophan residues are buried deeply and near to the surface, while in staphylococcal nuclease and glucagon, the tryptophan residues are exposed partially and fully to solvent. As the solvent accessibility of tryptophan is increasing, the emission spectra are shifted to a longer wavelength, shown by spectra 1 to 4 (Fig. 3.37 B).

Similarly, if tryptophan residues move away from the surface into the apolar environment a shift of emission spectra takes place towards a shorter wavelength (blue-shift) and if they became more open towards solvent the shift of emission spectra takes place toward a longer wavelength (red-shift)^{243,244}.

Quenching of the tryptophan fluorescence intensity is observed with an increasing gradient of phenolic compounds, shown in Fig. 3.38. The decrease in the tryptophan fluorescence intensity of NhGH11 is correlated with the concentration of phenolic compounds, indicating that a concentration-dependent quenching of the protein fluorescence intensity occurred. The quenching of the fluorescence intensity of NhGH11 can be explained by the Stern-Volmer fluorescence quenching mechanism²⁴⁵.

1. Photoexcitation: NhGH11 → NhGH11*

The excitation of NhGH11 was performed at 280 nm wavelength, which is an absorption maximum of tryptophan residues. The tryptophan residues acted as fluorophores; they absorbed photons and attained the excited state²⁴⁶.

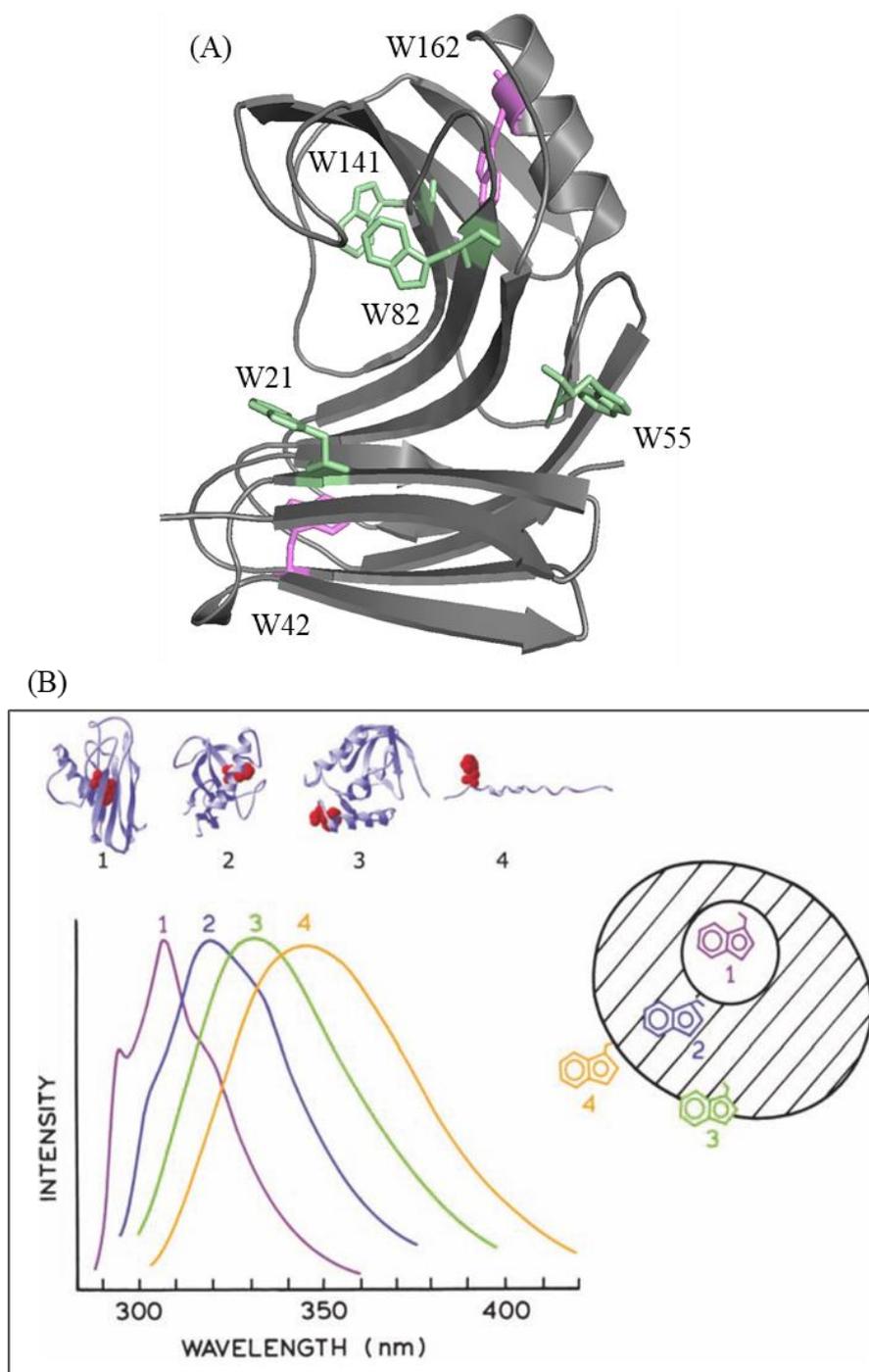


Fig. 3.37 Effect of tryptophan positions on emission spectra. (A) Cartoon representation of NhGH11 showing surface exposed (green) and buried tryptophan residues (purple). (B) The tryptophan emission spectra are shown with the solvent accessibility of tryptophan residues. 1. Apoazurin structure with deeply buried tryptophan, 2. Ribonuclease structure with tryptophan buried near to the surface, 3. Staphylococcal nuclease structure with partially solvent-exposed tryptophan. 4. Glucagon structure with solvent-exposed tryptophan²⁴³.

2. Fluorescence: NhGH11* → NhGH11

The fluorescence emission spectra of NhGH11 were observed between 320-420 nm. The excited electrons returned to the ground state by releasing the absorbed energy in the form of ultraviolet radiation which was observed as fluorescence emission spectra (control spectra in Fig 3.38).

3. Quenching: NhGH11* + Ph → NhGH11 + Ph*

Quenching of fluorescence occurred in the presence of phenolic compounds (Ph). The proposed explanation for quenching includes that the phenolic compounds acted as an electron acceptor and they quench the excited electrons from NhGH11. Therefore, quenching of fluorescence takes place from 10 to 1000 μM concentration of all phenolic compounds shown previously (Fig 3.38).

According to the NhGH11 spectrum data, lignin-derived phenolic compounds shifted the λ_{max} towards a higher wavelength, thus redshift took place, which correspond to the decrease of the emitted radiation in terms of frequency and photon energy²⁴⁸. The redshift promoted by the increasing concentrations of phenolic compounds took place after concentrations higher than 10 μM . Moreover, the decrease in fluorescence intensity was also observed in concentrations higher than 10 μM and may indicate the decrease of the excited state lifetime of the tryptophan residues.

The redshift added 10 nm emission spectra for caffeic and ferulic acid, while for coumaric and cinnamic acid less increments of emission spectra (6 nm and 2 nm) were observed. Redshift indicated that the binding of phenolic compounds resulted in solvent accessibility of tryptophan residues²⁴⁷, therefore, in the presence of caffeic and ferulic acid, the tryptophan residues of NhGH11 became more solvent-exposed as compared to coumaric and cinnamic acid, in other words, the binding of phenolic compounds with the higher number of hydroxyls can induce structural changes around the fluorophores (tryptophan residues) present in the structure of NhGH11.

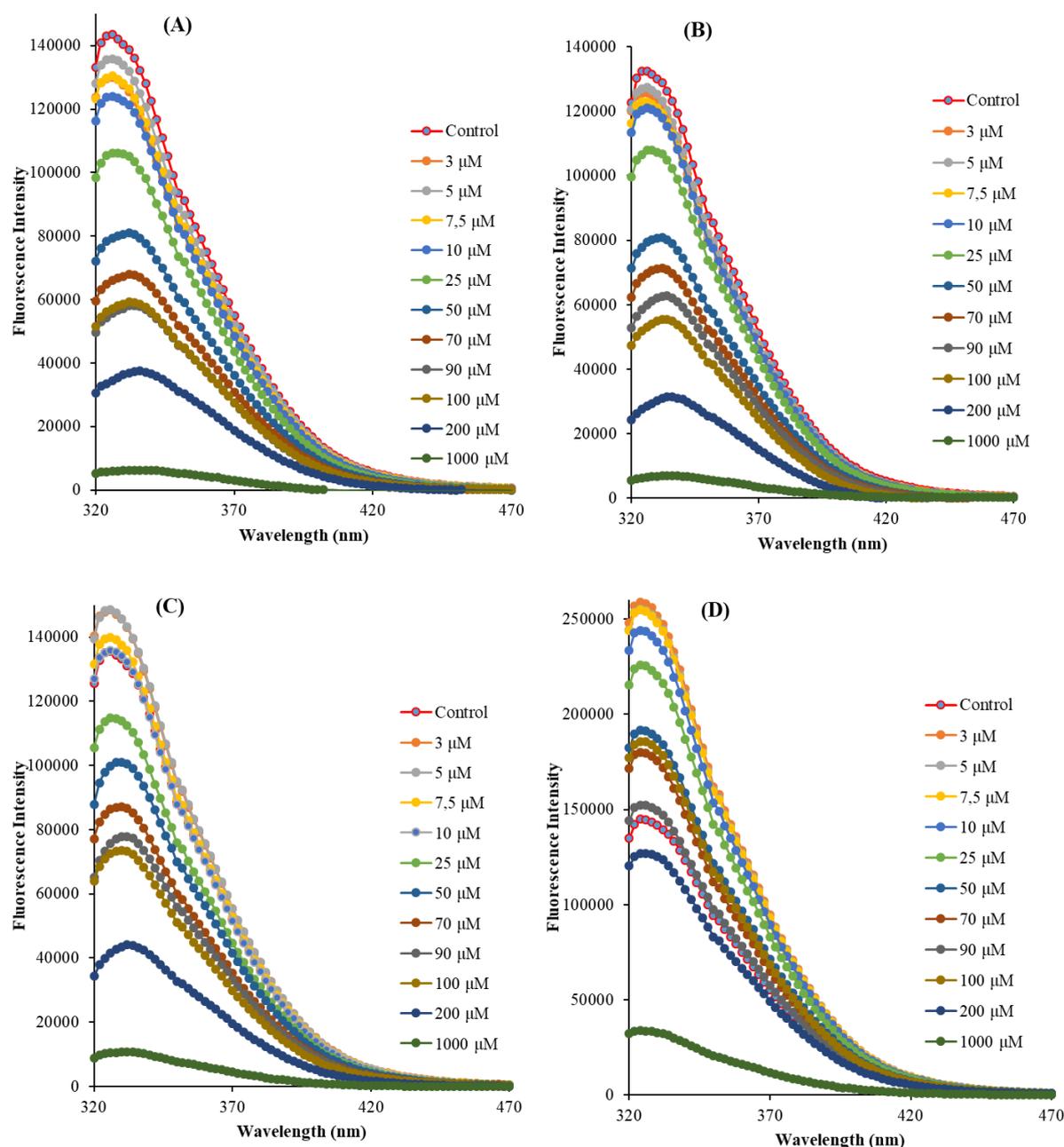


Fig. 3.38. Tryptophan fluorescence intensity of NhGH11. Recorded spectra of NhGH11 as control and NhGH11 with increasing concentrations of phenolic compounds from 3 up to 1000 μM . Four different compounds were applied: (A) caffeic acid, (B) ferulic acid, (C) coumaric acid and (D) cinnamic acid.

Non-linear regression curves (Fig. 3.39 A) were analyzed to calculate the binding constants of phenolic compounds. Caffeic acid showed binding at lower concentration ($80.79 \pm 6.6 \mu\text{M}$), thus it has higher binding affinity followed by ferulic acid ($100.2 \pm 7.4 \mu\text{M}$), coumaric acid ($124.5 \pm 11.8 \mu\text{M}$) and cinnamic acid ($183.7 \pm 8.9 \mu\text{M}$) as shown in Fig. 3.39 B.

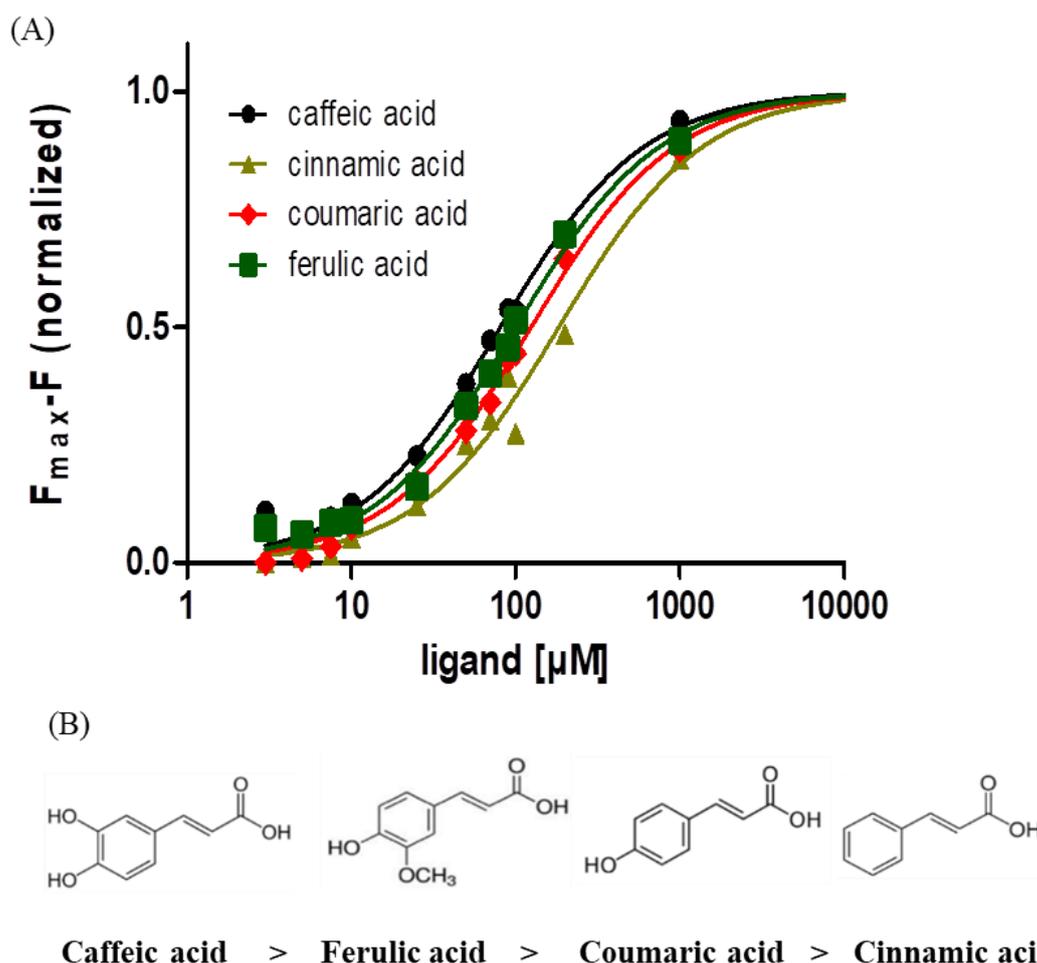


Fig. 3.39 Binding constants of phenolic compounds. (A) Non-linear regression curves for the calculations of binding constants of lignin-derived inhibitors. To determine the dissociation constants, the maximum of the emitted fluorescence was determined and plotted against the increasing concentration of phenolic compounds. (B) The binding tendency of phenolic compounds, caffeic acid showed higher binding followed by ferulic acid, coumaric acid and cinnamic acid.

Results of NhGH11 corroborated with GH11 xylanase from *T. xylanilyticus* that displayed a higher binding affinity for caffeic acid followed by coumaric acid, ferulic acid and cinnamic acid. In the same study, the binding constants of the phenolic compounds were calculated and were found to be depending on the number of hydroxyl groups of inhibitors. Therefore, the binding efficiency of these inhibitors was increased by increasing the number of hydroxyl groups on phenolic compounds²³¹.

A combination of all three assays provided evidence that these phenolic compounds affect the activity and the three dimensional structure of NhGH11. The binding pockets for phenolic compounds were predicted to study insight into the molecular basis of NhGH11 interactions with inhibitors, however, the uncertainty of binding pockets and the absence

of any crystallographic xylanase complex with these phenolic compounds obstructed the reliable docking studies. Whereas among the phenolic compounds analyzed, cinnamic acid showed a non-significant redshift and lower tendency of binding to NhGH11 (section 3.9.3), which may correlate with less possible hydrogen bond interactions due to the absence of hydroxyl groups.

Therefore, after biomass pretreatment, the nature and the concentrations of lignin-derived phenolic compounds present in the saccharide liquor must be characterized for better catalytic efficiency of enzymes used in biofuel industries.

3.10 Conclusion and outlook

Enzymatic hydrolysis of lignocellulosic waste material is cost effective and an alternative process for biofuel production, supporting the replacement of fossil fuels. Therefore, the research activities were focused on investigating the three-dimensional structure of xylanase from *Nectria haematococca* (NhGH11) as well as characterizing its functionality. The investigations and obtained results can be summarized as:

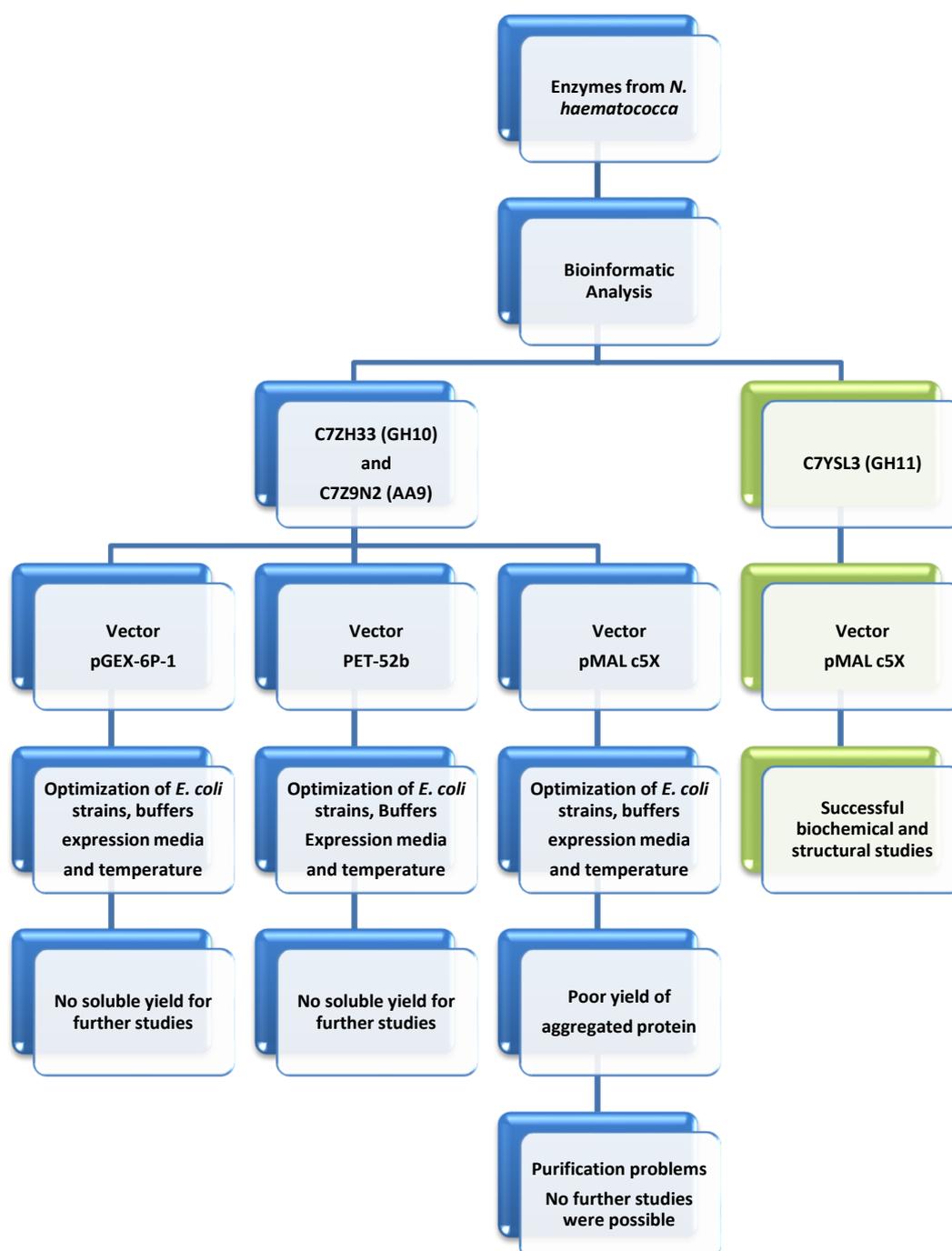
- *Nectria haematococca* xylanase (NhGH11) is the first enzyme studied from this fungus. The three-dimensional structure of NhGH11 shows the typical fold of family 11 xylanases with a single catalytic domain and conserved active site residues. The overall structure resembles a partially closed right hand, consisting of two major β -sheets and one α -helix. The crystal structure refined to 1.0 Å resolution, allows the investigation of hydrogen atom positions, alternative conformations of amino acid side chains and solvent molecules. The assignment of putative active site residues and their function in NhGH11 are proposed after superimposing and comparing the structure of the *T. reesei* (PDB code: 4HK8) enzyme-xylohexose complex, indicating the pair of glutamate 89 and glutamate 180 to be responsible for the catalytic reaction. Further, NhGH11 can accommodate a linear xylan backbone with six subsites.
- Structural and biochemical investigations provide an understanding of the molecular structure and activity parameters, such as optimum temperature, pH and thermostability. NhGH11 showed 100% activity at 45 °C and has also good activities at 40 °C, 45 °C and 50 °C, with approximate activities scaled to 92%, 100% and 84%. The structural basis of thermostability was investigated and monitored applying CD spectroscopy, indicating that the α -helix remains stable in an aqueous solution up to 45 °C and β -sheets remain partially stable up to 90 °C. The structural comparison of thermostability parameters with hyperthermostable xylanases showed the mesophilic character of NhGH11.
- NhGH11 is alkaline xylanase and the alkaline nature of NhGH11 is also observed during structural comparison with GH11 xylanases. The structure comparison showed that acidic xylanases have an aspartate residue, while alkaline xylanases and NhGH11 have asparagine residues vicinal to the catalytic glutamate at hydrogen bond distance.
- NhGH11 shows β -1,4 xylanase activity towards xylan substrates i.e. xylan

beechwood, xylan polysaccharide and wheat arabinoxylan. However, the specificity constant is found to be higher for xylan beechwood as compared to wheat arabinoxylan and xylan polysaccharides. Furthermore, NhGH11 also cleaved xylopentose substrates into small oligosaccharides i.e. xylotetrose, xylotriose, xylobiose and xylose, analyzed by MALDI-TOF-MS (matrix-assisted laser desorption ionization-time of flight mass spectrometry).

- The role of different metal ions and chemical reagents was studied for reactions catalyzed by NhGH11. As a result, the data showed that NhGH11 is relatively less sensitive towards most common metal ions and chemical reagents.
- Lignin derived phenolic compounds decreased the enzymatic activity of NhGH11 at concentrations higher than 1 mM. Calculation of melting temperature in presence of phenolic compounds showed a concentration-dependent destabilization of NhGH11 structure. However, cinnamic acid showed stronger effect followed by coumaric, ferulic and caffeic acid. The possible interactions of phenolic compounds with NhGH11 can be explained by a comparison of possible hydrogen bond interactions. The structure of NhGH11 has potential binding pockets that can form hydrogen bonds. Caffeic acid can form more hydrogen bonds followed by ferulic, coumaric and cinnamic acid depending on the number of hydroxyl groups. The calculated binding constants also provided evidence that caffeic acid has more binding capacity followed by ferulic, coumaric and cinnamic acid.

The structural information of NhGH11 can be used to understand the molecular basis of enzymatic specificity and highlighting its functional features to be applied in industrial processes. Also, biochemical studies indicate the potential of NhGH11 for commercial applications. NhGH11 can be applied in industrial processes requiring β -1,4 xylanase activity at working temperature ranging from 40 to 50 °C, and requiring a pH range from 5.5-6.0. Also, obtained xylooligosaccharide products can be also used in food industries. In this context, NhGH11 can be used to supplement enzymatic cocktails used for biofuel production. However, future research activities can explore all glycosyl hydrolases from *Nectria haematococca* to be applied in formulations of enzymatic cocktails used in biofuel industries.

4. Supplementary information



S. 4.1 Flowsheet diagram. Three different families of biofuel enzymes were selected from *N. haematococca*. *E. coli* expression of GH10 and AA9 did not yield soluble protein, while GH11 xylanase (NhGH11) is successfully applied for structural and biochemical studies.

The *E. coli* expression of GH10 and AA9 from *N. haematococca* resulted in the formation of inclusion bodies. Experiments were performed for optimization by using different expression vectors, buffers, expression temperatures and *E. coli* strains, as shown in S. 4.1. However, soluble fractions of these enzymes could not be obtained for further biochemical and biophysical characterizations. These solubility problems are correlated to the findings of other *N. haematococca* enzymes where expression problems by *E. coli* also resulted in the extensive formation of inclusion bodies²⁴⁹. Although, *E. coli* is a preferable host for cost-effective recombinant expression of industrially important enzymes, due to inexpensive growth conditions, an easy transformation of an external gene and higher expression yields. However, many fungal enzymes cannot be expressed by *E. coli* or resulted in the formation of inclusion bodies²⁵⁰. The reasons for *E. coli* expression problems for eukaryotic proteins may be related to the repetition of rare codons or the absence of post-translational modifications such as glycosylation²⁵¹.

S. 4.2 Alignment of GH11 sequences

	1	10	20
C7YSL3*	MV	SFSYLL.AAVSAFTSVL	AVPT
A2R5J7	MV	SFLGQARL.AVPILSAF	ACMLAASSA
Q2UFR7	MV	S.ATRLLL.LLPFLGAL	ASPTD.PTP
Q2TYR4	MV	S.ASQLLF.CLPVLGAL	ATPTD.SIT
Q9AG99			
Q96TR7	MKFF	FATIAALVVAA	VAPVAEADA
Q12562	MKFF	FATIAALVVGA	VAPVAEAEA
Q9UW17	MKFF	FATIAALVVAA	VAPVAEAEA
A2I7V2	MKV	TAAFAALLVTA	FASVP
Q12549	MKV	TAAFASLLLTA	FAPAP
E3UN71	MKV	TAAFASLLLTA	FAPAP
A6N2L8			
Q45UD8	MKV	TAAFASLLLTA	FAPAP
Q2I0I8	MKV	TAAFAGLLVTA	LAPAP
Q12550	MKV	TAAFAGLLVTA	FAPVP
P55331	MKV	TAAFAGLLVTA	FAPAP
P55328	MKV	TAAFAGLLVTA	FAPVP
P33557	MKV	TAAFAGLLVTA	FAPVP
P55329	MKV	TAAFAGLLVTA	FAPVP
Q6U894	MF	KFVKVLTVVI	AATISFC
Q71S35	MF	KFVKVLTVVI	AATISFC
Q43993	MF	KFGKKLMTVVV	AASMSFG
Q1XGE6	MF	KFGKKLLTVV	AASMSFG
Q6TLP3	MF	KLKILMLVLL	TISMFSIS
B5M6I0	MF	KFKKNFLV.GL	TAALMSIS
E0YL13	MF	KFKKKFLV.GL	TAAFMSIS
A2T401	MF	KFKKNFLV.GL	TAALMSIS
C5MTD6	MF	KFKKNFLV.GL	SAALMSIS
Q3HLJ4	MF	KFKKNFLV.GL	SAALMSIS
Q8RMN8	MF	KFKKNFLV.GL	SAALMRII
Q45VU6	MF	KFKKNFLV.GL	SAALMSIS
Q45VU3	MF	KFKKNFLV.GL	SAALMSIS
P09850	MF	KFKKNFLV.GL	SAALMSIS
P18429	MF	KFKKNFLV.GL	SAALMSIS
Q9RMM4	MN	TL	VHPQGRAGGLR
D7EZJ3	MHDAP	AQRKRRRPRGRIGPLPRSSRFAR	LK.LLIASACAALLATL
C6WIK2		MAEAADAAAGGQHRIGR	R.LFLGGAGATVVGAAAGL
Q8GMV7	MNEPL	TITQARRRRRLGLRRIVT	S.AF
Q9RI72	MQQDGT	QDRIKQSPAPLNGMSR	R.GFLGGAGTLALATASGL
Q56013	MQQD	GKRQDQNPAPFSGLSR	R.GFLGGAGTVALATASGL
Q59962		MTKDNTPIRPVSR	R.FIGRAGALAL.ATSGL
P26515		MNLLVQPRRRRRRGPVILLVR	S.AWA
B2LWN3			MISLKRVAALLCVAGLGMS
B3PIN0			MKIFQNTKNVIVSIAWAAALCT
C5BU24			MKLL.SLRNTLSALALAI
Q92244			Q.ILTWAL.AALAAIPAVT
O43097			MV.GFTPVA.LAALAAATGAL
Q0ZBL0			MV.NFSSLL.LAASAAVGAV
Q00263			MV.SFSSIF.TAAVAATGAL
O74716			MF.LTSVVSLVVGAI
Q9HFH0			MK.LFLAAIVLCATAATAFP
A2QBA9			
E0X4B3			MV.SFSS.LVLAASTVAGVL
P87037			MV.SFSS.LLLAVSAVSGAL
Q0Q592			MV.AFSSNLLALG.SATAAL
B0FX61			MLGSILVALT.AAAGAI
Q0ZHI9			MV.SFKSLLLAAS.AFTAVL
Q0ZBK9			MV.SFKALLFGATGALALP
D3KT79			MA.RLSQFL.LTCALAVKAF
P48824			ML.T.KNLLLCFAAAKAVL
Q6QA21			ML.T.KNLLLCFAAAKAVL
F5CI28			ML.T.KNLLLCFAAAKAVL
Q3S401			ML.T.KNLLLCFAAAKAVL
A2Q7I0			ML.T.KNLLLCFAAAKAAL
P55330			ML.T.KNLLLCFAAAKAAL
Q2PU02			ML.T.KNLLLCFAAAKAAL
Q9HFA4			MV.SFSSILLACSAA.IGAL
P55333			MV.SFSSL.LLACSA.VTAF
P55332			MV.SFKSLLVLCCLAA.LGAF
A1CCU0			MV.SFKYL.FLAASA.LGAL
Q4WG11			MV.SFSYL.LLACSA.IGAL
Q69IG4			
Q6QHA0			MKNWPGTGIILLLAGGLLA
D4PAK8			MGEQV
A4UWQ6			MKFVFAF
E2JFA2			MAAML
P77853			MF.LKLL
P17137			MLRRKVIFIVLATLVMTSLTIVDNT
Q7SID8			
E2IHA1			MDLRK
Q8RMN7			MNLRK
B5MOC6			MNLRK

4. Supplementary information

			β_1	β_2
C7YSL3*	...AETS.....	EN.....	PALSKRTQSSTG....	THGGY YYSF
A2R5J7	IPPPPRGALSPERLQWIREVIGNQTENDSVSDLAKRSTILHTSQDGV....		DSAGF	YYSV
Q2UFR7	NEPVGGRGKESYDLLQEIGESIFNAISE...DKLQGRSTDLRTSKDGV....		NSAGY	YYSL
Q2TYR4	AQAVARS...PEFLEHMGALIANATGE...ATLERRDATFKTSKDGV....		DAAGF	YYSL
Q9AG99			
Q96TR7	E...AS.....	S.....	PMLIE...RAG....	PGGIN YVQN
Q12562	E...AS.....	S.....	PMLIE...RAG....	PGGIN YVQN
Q9UW17	E...AS.....	S.....	PMLIE...RAG....	PGGIN YVQN
A2I7V2	E.....	PVLV...SR....	SAGIN YVQN
Q12549	E.....	PVLV...SR....	SAGIN YVQN
E3UN71	E.....	PVLV...SR....	SAGIN YVQN
A6N2L8			SAGIN YVQN
Q45UD8	E.....	PVLV...SR....	SAGIN YVQN
Q2I0I8	G.....	PVLV...SR....	SAGIN YVQN
Q12550	E.....	PVLV...SR....	SAGIN YVQN
P55331	E.....	PDLV...SR....	SAGIN YVQN
P55328	E.....	PVLV...SR....	SAGIN YVQN
P33557	E.....	PVLV...SR....	SAGIN YVQN
P55329	E.....	PVLV...SR....	SAGIN YVQN
Q6U894		LSAVP.....	ASANT YWQY
Q71S35		LSAVP.....	ASANT YWQY
Q43993		VFAA.TS....	SAATD YWQN
Q1XGE6		VFAA.TS....	SAATD YWQN
Q6TLP3		LFTVTAY....	AASTD YWQN
B5M6I0		LFSASAS....	AAGTD YWQN
E0YL13		MFSATAS....	AAGTD YWQN
A2T401		LFSATAS....	AASTD YWQN
C5MTD6		LFSATAS....	AASTD YWQN
Q3HLJ4		LFSATAS....	AASTD YWQN
Q8RMN8		LFSATAS....	AASTD YWQN
Q45VU6		LFSATAS....	VASTD YWQN
Q45VU3		LFSATAS....	AASTD YWQN
P09850		LFSATAS....	AASTD YWQN
P18429		LFSATAS....	AASTD YWQN
Q9RMM4		TARAD.TITSNQTG....	THNGY FYSF
D7EZJ3		AAHAQ.TVTSNQTG....	NHNGY FYSF
C6WIK2		VAHAQ.TITTNQTG....	THNGY FYSF
Q8GMV7		TAHADTTITQNTG....	YDNGY FYSF
Q9RI72		TAHAATTTITNTGTG....	T.DGM YYSF
Q56013		TAHAATTTITNTGTG....	Y.DGM YYSF
Q59962		TARADTVITNTGTG....	TNNGY YYSF
P26515		TAQADTVVTTNQTG....	TNNGY YYSF
B2LWN3		AANAQTCLTSSQTG....	TNNGF YYSF
B3PIN0		SAVSAQTLTNSSTG....	TNNGF YYSF
C5BU24		AVNAQQTLTNSSTG....	THGGH YYSF
Q92244	S S M D E L V	E R . S	P N V T L V A R G T P S S T G	T H N G F Y Y S H
O43097	...GNAT.....	E.....	LEKRQTTTPNSEG....	WHDGY YYSW
Q0ZBL0	...ELPG.....		MNINKRQTYTSSATG....	THNGY YYSF
Q00263	...TDLA.....	T R . S	L G A L T A R A G T P S S Q G	T H N G C F Y S W
O74716		SELMQMTPRNSC....	YGGGL YYSY
Q9HFH0		GDLSKRQSITTSQTG....	TNNGY YYSF
A2QBA9			
E0X4B3		VEL.AKRQLTSSQTG....	TNNGY YYSF
P87037		VEL.AKRAITSSSETG....	TNNGY YYSF
Q0Q592	AQAQEE.....		SANLTARATPAGTG....	MNNGF FYSF
B0FX61		DLITRSTPAGTG....	TNNGF FYSF
Q0ZHI9	PD...VN.....	I T D A	D E L L V R R Q V T A N S E G	T H N G Y F Y S W
Q0ZBK9	REAHARG.....	E N . V	T E L L M A R A G T P S Q T G	W N N G Y Y Y S F
D3KT79	V...EER.....	G P . N	F F S A L T E R S T G S S T G	Y S N G Y Y Y S F
P48824	...VE.....	R S . D	A L H K L S E R S T P S S T G	E N N G Y Y Y S F
Q6QA21	...VE.....	R S . D	A L H K L S E R S T P S S T G	E N N G F Y Y S F
F5CI28	...VE.....	R S . D	A L H K L S E R S T P S S T G	E N N G F Y Y S F
Q3S401	...VE.....	R S . D	A L H K L S E R S T P S S T G	E N N G F Y Y S F
A2Q7I0	...AQ.....	R S . D	A L H M L S E R S T P S S T G	E N N G F Y Y S F
P55330	...AQ.....	R S . D	A L H M L S E R S T P S S T G	E N N G F Y Y S F
Q2PU02	...AQ.....	R S . D	A L H M L S E R S T P S S T G	E N N G F Y Y S F
Q9HFA4	HPNEAFN.....	E T . A	F N D . L V G R S T P S S T G	Y N N G Y Y Y S F
P55333		AERLSERSTPSSTG....	TSGGY YYSF
P55332	...AARE.....	A S	L L E R S T P S S T G	W S N G Y Y Y S F
A1CCU0	...SWFN.....	E T . A	L H E F A E R A G T P S S T G	W N N G Y Y Y S F
Q4WG11	...SFN.....	E T . A	L H E F A E R A G T P S S T G	W N N G Y Y Y S F
Q69IG4	AKA.....	Q W G G G	A S A G Q K L S V G G G Q N Q H K G V S D G F	S Y E I
Q6QHA0	GKR.....	Q E G Q K	V D P D T Q N E K L T G G T V F T A N S R G N R P L E G S P Y G Y E M	G Y E M
D4PAK8	VKP.....	T D E R I	I D P S T A N T Q L T G G V T Y N T T S G G N K P L A G S P Y G Y E T	G Y E T
A4UWQ6	.AL.....	A F A G S	F E P L V N N T A S N A E C T V T S N S Q G	N C G G V A Y E L
E2JFA2	GKP.....	A D S K S S E S K T . T E A E K E V E H S V E A Q E F N E N I I G	T A D G Y G Y E L	G Y E L
P77853	S K L L L V L L V A V Y T . Q V N A Q T S I T L T S N A S G	T F D G Y Y Y E L	Y Y E L
P17137	TES.....	T F S K E V L S T . Q K T Y S A F N T Q A A P K T I T S N E I G	V N G G Y D Y E L	Y E L
Q7SID8		ATTITSNQTG....	THDGY D Y E L
E2IHA1	LRL.....	L F . . . V M C I . G L T L I L T A V P A H A R T I T N N E M G	N H S G Y D Y E L	Y E L
Q8RMN7	LRL.....	L F . . . V M C I . G L T L I L T A V P A H A R T I T N N E M G	N H S G Y D Y E L	Y E L
B5MOC6	LRL.....	L F . . . V M C I . G L T L I L T A V P A H A R T I T N N E M G	N H S G Y D Y E L	Y E L

4. Supplementary information

	β^2		β^3		β^4		β^5				
	50	71 922	60	70	80	90					
C7YSL3*	WIDN	PNTV	TYT	...NQNA	GQFSV	SW	...GN	QGNFV	GKGWNP	GAR	
A2R5J7	YNDNG	ADV	GYT	...EYPTT	GQFEL	GS	...AE	AEFLA	GKGFKG	GNPR	
Q2UFR7	YNDNH	AGA	DYT	...EFPDS	GRFQL	KN	...TN	KEFLG	GKGYRG	GS	
Q2TYR4	YNANG	AGA	EYS	...ESENS	GQFKLS	WN	...TN	SEFLG	GKGFKG	GS	
Q9AG99											
Q96TR7	YNGN	LQF	TYN	...ENA	GTYSM	WT	...NGVS	GDFVV	GLGWS	TGAR	
Q12562	YNGN	LQF	TYN	...ENA	GTYSM	WN	...NGVN	GDFVV	GLGWS	TGAR	
Q9UW17	YNGN	LQF	TYN	...ENA	GTYSM	WN	...NGVN	GDFVV	GLGWS	TGAR	
A2I7V2	YNGN	LQF	TYD	...ESA	GTYSM	WE	...DGVS	SDFV	DL	...TFVS	
Q12549	YNGN	LQF	TYD	...EST	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
E3UN71	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
A6N2L8	YNGN	LQF	RYD	...EST	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
Q45UD8	YNGN	LQF	TYD	...EST	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
Q2I0I8	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
Q12550	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
P55331	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
P55328	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
P33557	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
P55329	YNGN	LQF	TYD	...ESA	GTFSM	WE	...DGVS	SDFV	GLGWT	TGS	
Q6U894	WIDGG	GTV	NAT	...NGPG	GNYSV	WR	...DT	GNFV	GKGW	EIPNR	
Q71L35	WIDGG	GTV	NAT	...NGPG	GNYSV	WR	...DT	GNFV	GKGW	EIPNR	
Q43993	WIDGG	GTV	NAV	...NGSG	GNYSV	WQ	...NT	GNFV	GKGW	YIPNR	
Q1XGE6	WIDGG	GTV	NAV	...NGSG	GNYSV	WQ	...NT	GNFV	GKGW	YIPNR	
Q6TLP3	WIDGG	GTV	NAT	...NGSD	GNYSV	WS	...NC	GNFV	GKGW	TGSA	
B5M6I0	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
E0YL13	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
A2T401	WIDGG	GMV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPSR	
C5MTD6	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPSR	
Q3HLJ4	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
Q8RNM8	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
Q45VU6	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
Q45VU3	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
P09850	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
P18429	WIDGG	GTV	NAV	...NGSG	GNYSV	WS	...NT	GNFV	GKGW	TGSPFR	
Q9RMM4	WIDAP	GTV	TMN	...TGAG	GNYST	WS	...NT	GNFV	GKGW	TGSR	
D7E2J3	WIDAP	GTV	SAT	...MGSG	GNYST	WR	...NT	GNFV	GKGW	TGSR	
C6WIK2	WIDSQ	NTV	SMT	...LGGG	GNYST	WR	...NT	GNFV	GKGW	TGSR	
Q8GMV7	WIDAP	GTV	SMT	...LHSG	GSYST	WR	...NT	GNFV	GKGW	TGSR	
Q9RI72	WIDGG	GTV	SMT	...LNGG	GSYST	WT	...NC	GNFV	GKGW	TGSR	
Q56013	WIDGG	GTV	SMT	...LNGG	GSYST	WT	...NC	GNFV	GKGW	TGSR	
Q59962	WIDGG	GTV	SMT	...LNGG	GSYST	WT	...NC	GNFV	GKGW	TGSR	
P26515	WIDSQ	GTV	SMN	...MGSG	GNYST	WR	...NT	GNFV	GKGW	TGSR	
B2LWN3	WKDSP	GTV	NFC	...LQSG	GRYTS	WS	...G	INN	WVGK	GKGW	TGSR
B3PIN0	WKDS	GDA	SMT	...LLSG	GRYQS	WN	...SS	TNN	WVGK	GKGW	TGSPSR
C5BU24	WKDS	GNA	SFT	...LYDG	GRYGS	WN	...SG	TNN	WVGK	GKGW	TGSPGA
Q92244	WIDNAG	ADV	TYS	...MGGG	GQFSV	WR	...NS	GNFV	GKGW	TGSPGN	
043097	WSDG	GAQA	TYT	...NLEG	GTYE	ISWG	...DG	GNLV	GKGW	TGSPGL	
Q0ZBL0	WIDG	QGS	QYT	...NEAG	GQYSA	RS	...GN	GNWV	GKGW	TGSPGA	
Q00263	WIDG	GAQA	TYT	...NGAG	GSYSV	WK	...TG	GNLV	GKGW	TGSPGA	
074716	WADYG	NT	RY	...CGAG	GHYDL	SWG	...NG	GNV	ARG	GKWPAS	PR
Q9HFH0	WINGG	GEV	TYT	...NGDN	GEYSV	TWV	...DC	GDF	TSGK	GKWN	PAN
A2QBA9				...NGDA	GEYTV	EW	...DC	GDF	VAGK	GKWN	PAS
E0X4B3	WIDGG	GTV	TYT	...NGNG	GQYQV	DN	...NC	GNFV	GKGW	TGSP	
P87037	WINGG	GTV	EYT	...NGNG	GQYSV	KWT	...NC	DNFV	GKGW	TGSP	
Q0Q592	WIDGG	GTV	TYN	...NGAA	GKYDV	QS	...NV	GNFV	GKGW	TGSP	
B0FX61	WIDGA	GTV	TYN	...NGDA	GSYDV	TWQ	...NV	GNFV	GKGW	TGSP	
Q0ZHI9	WSDGG	GTV	TYT	...MGAG	SRYSV	TWK	...DT	GNFV	GKGW	TGSP	
Q0ZBK9	WIDNG	GTV	NYW	...NSAN	GGYGI	NS	...NT	GNFV	GKGW	TGSP	
D3KT79	WIDGAS	GTV	EYS	...NGAG	GSYSV	TWS	...SA	SNFV	GKGW	TGSP	
P48824	WIDGG	GTV	TYT	...NGNA	GSYSV	EWS	...NV	GNFV	GKGW	TGSP	
Q6QA21	WIDGG	GTV	TYT	...NGDA	GSYIV	EWS	...NV	GNFV	GKGW	TGSP	
F5CI28	WIDGG	GTV	TYT	...NGDA	GSYIV	EWS	...NV	GNFV	GKGW	TGSP	
Q3S401	WIDGG	GTV	TYT	...NGDA	GSYIV	EWS	...NV	GNFV	GKGW	TGSP	
A2Q7I0	WIDGG	GTV	TYT	...NGDA	GAYIV	EWS	...NV	GNFV	GKGW	TGSP	
P55330	WIDGG	GTV	TYT	...NGDA	GAYIV	EWS	...NV	GNFV	GKGW	TGSP	
Q2PU02	WIDGG	GTV	TYT	...NGDA	GAYIV	EWS	...NV	GNFV	GKGW	TGSP	
Q9HFA4	WIDGG	GTV	TYT	...NGNG	GSYSV	QWS	...NV	GNFV	GKGW	TGSP	
P55333	WIDGG	GTV	TYT	...NGDG	GSYIV	EW	...NV	GNFV	GKGW	TGSP	
P55332	WIDGG	GTV	TYT	...NGAG	GSYSV	QWS	...NV	GNFV	GKGW	TGSP	
A1CCU0	WIDNG	GTV	NYQ	...NGNG	GSYSV	QWK	...DT	GNFV	GKGW	TGSP	
Q4WG11	WIDGG	GTV	TYT	...NGAG	GSYSV	WR	...NV	GNFV	GKGW	TGSP	
Q69IG4	WLDNTG	GSG	SMILG	...SGAT	FKAE	WNA	...NR	GNFL	ARRGL	DFGS	QK
Q6QHA0	WTQ	GGNN	NKLV	WF	GGPD	QGGG	AA	FRA	EW	...EP	DD
D4PAK8	WIDTGG	GVC	SLTW	YGAD	QGGG	AA	FKA	TWT	...NP	HDFL	GR
A4UWQ6	WMSG	SGSG	C	TFKGD	...GSA	AFNA	QWN	...NC	GDYL	C	RAG
E2JFA2	WKD	NGDI	...KFF	V	GG	GC	F	S	CE	WK	...NI
P77853	WKD	TGNT	...TMT	VY	...TQ	GR	F	S	C	QWS	...NI
P17137	WKDY	GN	...S	MLK	...NG	GA	F	S	C	QWS	...NI
Q7SID8	WKD	SGNT	...S	MLN	...SG	GA	F	S	C	QWS	...NI
E2IHA1	WKDY	GN	...S	MLN	...NG	GA	F	S	C	QWS	...NI
Q8RNM7	WEDY	GN	...S	MLN	...NG	GA	F	S	C	QWS	...NI
B5MOC6	WKDY	GN	...S	MLN	...NG	GA	F	S	C	QWS	...NI

4. Supplementary information

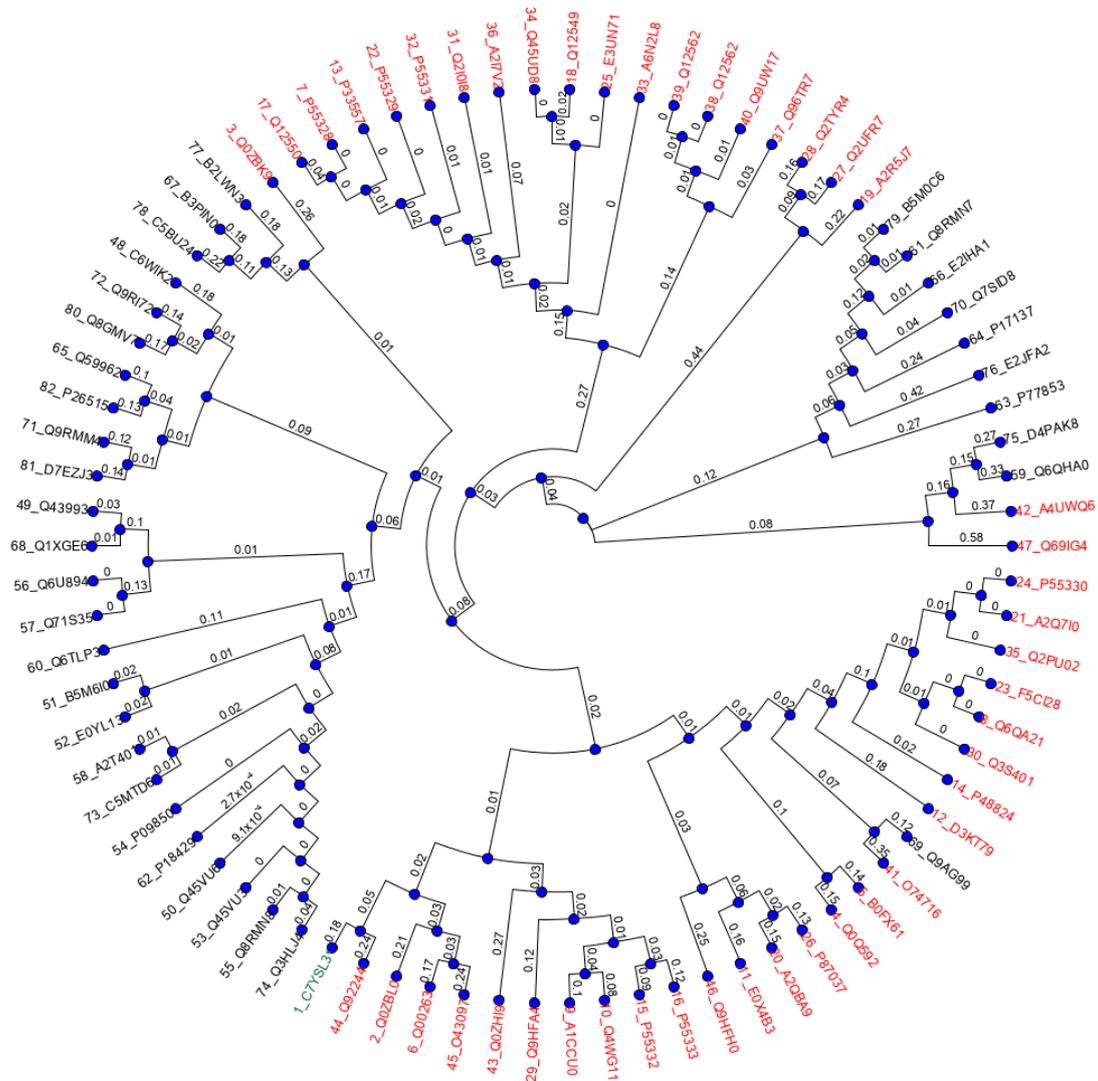
	β_6			β_7			β_8																																		
	100			110			120																																		
C7YSL3*	T	I	K	Y	S	G	T	Y	N	P	N	G	N	S	Y	L	A	V	Y	G	W	T	R	N	P	L	I	E	Y	I	V	E	N	F	G	T	Y	N			
A2R5J7	S	L	T	W	D	G	Y	F	T	A	E	G	D	W	T	L	A	I	Y	G	W	T	I	N	P	V	T	E	W	Y	I	V	E	S	H	G	S	G	T		
Q2UFR7	K	L	T	W	D	G	H	F	K	A	D	G	D	Y	T	L	A	V	Y	G	W	T	I	D	P	V	T	E	W	Y	V	V	E	Q	H	G	T	G	T		
Q2TYR4	S	L	S	W	D	G	Q	F	Q	A	E	G	D	D	F	T	L	A	V	Y	G	W	T	I	D	P	V	T	E	W	Y	V	V	E	A	H	G	T	G	T	
Q9AG99																																									
Q96TR7	S	I	T	Y	S	S	S	Y	T	A	S	G	S	G	S	Y	L	S	V	Y	G	W	I	N	S	P	Q	A	E	Y	I	V	E	S	Y	G	S	Y	N		
Q12562	S	I	T	Y	S	S	N	Y	Q	A	S	G	S	G	S	Y	L	S	V	Y	G	W	I	N	S	P	Q	A	E	Y	I	V	E	S	Y	G	S	Y	N		
Q9UW17	S	I	T	Y	S	S	N	Y	Q	A	S	G	S	G	S	Y	L	S	V	Y	G	W	I	N	S	P	Q	A	E	Y	I	V	E	S	Y	G	S	Y	N		
A2I7V2	S	I	T	F	S	A	D	Y	S	A	S	G	S	G	S	Y	L	A	V	Y	G	W	V	N	S	P	Q	A	E	Y	I	V	D	K	Y	G	N	Y	N		
Q12549	S	I	T	Y	S	A	Q	Y	S	A	S	S	S	S	S	Y	L	A	V	Y	G	W	V	N	S	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
E3UN71	S	I	T	Y	S	A	Q	Y	S	A	S	S	S	S	S	Y	L	A	V	Y	G	W	V	N	S	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
A6N2L8	S	I	T	Y	S	A	Q	Y	S	A	S	S	S	S	S	Y	L	A	V	Y	G	W	V	N	S	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
Q45UD8	S	I	T	Y	S	A	Q	Y	S	A	S	S	S	S	S	Y	L	A	V	Y	G	W	V	N	S	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
Q2I0I8	P	I	T	Y	S	A	D	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	Y	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
Q12550	A	I	T	Y	S	A	E	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	L	S	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
P55331	A	I	T	Y	S	A	E	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	Y	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
P55328	A	I	T	Y	S	A	E	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	Y	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
P33557	A	I	T	Y	S	A	E	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	Y	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
P55329	A	I	T	Y	S	A	E	Y	S	A	S	G	S	S	S	Y	L	A	V	Y	G	W	V	N	Y	P	Q	A	E	Y	I	V	E	D	Y	G	D	Y	N		
Q6U894	T	I	H	Y	N	A	G	V	W	E	P	S	G	N	G	Y	L	T	L	Y	G	W	T	R	N	Q	L	I	E	Y	V	V	D	N	W	G	T	Y	R		
Q71535	T	I	H	Y	N	A	G	V	W	E	P	S	G	N	G	Y	L	T	L	Y	G	W	T	R	N	Q	L	I	E	Y	V	V	D	N	W	G	T	Y	R		
Q43993	V	V	N	Y	N	A	G	V	F	A	P	S	G	N	G	Y	L	T	F	Y	G	W	T	R	N	A	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
Q1XGE6	V	V	N	Y	N	A	G	V	F	S	P	S	G	N	G	Y	L	T	F	Y	G	W	T	R	N	A	L	I	E	Y	V	V	D	N	W	G	T	Y	R		
Q6TLP3	V	I	N	Y	N	A	G	A	F	S	P	S	G	N	G	Y	L	A	L	Y	G	W	T	R	N	S	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
B5M6I0	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
E0YL13	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	A	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
A2T401	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
C5MTD6	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
Q3HLJ4	T	I	N	Y	N	A	G	V	W	A	P	N	G	D	G	Y	L	T	L	Y	G	W	T	R	S	P	L	H	R	I	L	C	V	D	S	W	G	T	Y	R	
Q8RMN8	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	P	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
Q45VU6	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
Q45VU3	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
P09850	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
P18429	T	I	N	Y	N	A	G	V	W	A	P	N	G	N	G	Y	L	T	L	Y	G	W	T	R	S	P	L	I	E	Y	V	V	D	S	W	G	T	Y	R		
Q9RMM4	T	V	I	Y	S	G	T	F	N	P	S	G	N	A	Y	L	A	L	Y	G	W	S	Q	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
D7E2J3	T	V	I	Y	S	G	S	F	N	P	S	G	N	A	Y	L	T	L	Y	G	W	S	R	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
C6WIK2	T	V	N	Y	S	G	S	F	N	P	S	G	N	G	Y	L	C	L	Y	G	W	T	S	N	P	L	V	E	Y	I	V	E	S	W	G	N	Y	R			
Q8GMV7	T	V	I	Y	S	N	A	S	F	N	P	S	G	N	A	Y	L	T	L	Y	G	W	T	R	N	P	L	V	E	Y	I	V	E	S	W	G	T	Y	R		
Q9RI72	T	V	R	Y	N	G	Y	F	N	P	S	G	N	G	Y	G	C	L	Y	G	W	T	S	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
Q56013	T	V	R	Y	S	G	Y	F	N	P	S	G	N	G	Y	G	C	L	Y	G	W	T	S	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
Q59962	T	V	N	Y	S	G	S	F	N	P	S	G	N	A	Y	L	T	L	Y	G	W	T	A	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
P26515	T	V	Q	Y	S	G	S	F	N	P	S	G	N	A	Y	L	A	L	Y	G	W	T	S	N	P	L	V	E	Y	I	V	D	N	W	G	T	Y	R			
B2LWN3	N	I	T	Y	S	G	S	F	N	P	S	G	N	G	Y	L	A	L	Y	G	W	T	I	N	P	L	V	E	Y	I	V	V	D	S	W	G	W	R			
B3P1N0	V	I	S	Y	S	G	Y	G	V	D	S	S	Q	N	S	Y	L	A	L	Y	G	W	T	R	S	P	L	I	E	Y	I	V	I	E	S	Y	G	S	Y	N	
C5BU24	V	V	N	Y	E	G	Y	G	V	N	S	Q	N	S	Y	L	A	L	Y	G	W	T	R	N	P	L	I	E	Y	I	I	E	S	Y	G	S	Y	N			
Q92244	V	I	N	Y	S	G	S	Y	S	P	Q	G	N	S	Y	L	A	V	Y	G	W	T	R	N	P	L	I	E	Y	I	V	V	E	S	F	G	S	Y	N		
Q43097	A	I	H	F	E	G	V	Y	Q	P	N	G	N	S	Y	L	A	V	Y	G	W	T	R	N	P	L	V	E	Y	I	V	E	N	F	G	T	Y	D			
Q0ZBL0	T	I	N	Y	T	A	T	Y	N	P	N	G	N	S	Y	L	A	V	Y	G	W	T	R	N	P	L	I	E	Y	I	V	V	E	N	F	G	T	Y	D		
Q00263	T	I	T	Y	S	G	T	Y	S	P	S	G	N	S	Y	L	A	V	Y	G	W	T	R	N	P	L	I	E	Y	I	V	V	E	N	F	G	T	Y	D		
Q74716	A	V	T	Y	S	G	S	W	Q	C	N	G	N	C	Y	L	S	V	Y	G	W	T	I	N	P	L	V	E	Y	I	V	E	N	G	Y	N	Y	N			
Q9HFH0	T	V	I	Y	S	G	S	F	N	P	S	G	N	A	Y	L	A	V	Y	G	W	T	I	D	P	L	V	E	Y	I	I	L	E	S	Y	T	Y	N			
A2QBA9	A	V	T	Y	S	G	S	W	E	T	D	A	N	A	Y	L	S	V	Y	G	W	T	S	P	L	V	E	F	Y	I	V	D	K	Y	G	D	Y	D			
E0X4B3	A	V	T	Y	S	G	S	W	Q	T	S	G	N	G	Y	L	S	V	Y	G	W	T	S	P	L	V	E	F	Y	I	V	E	S	Y	G	D	Y	D			
P87037	T	V	I	Y	S	G	E	W	E	S	N	S	N	S	Y	V	S	L	Y	G	W	T	Q	N	P	L	V	E	Y	I	V	D	K	Y	G	D	Y	D			
Q00592	V	V	T	Y	S	G	T	W	N	A	A	N	V	N	S	Y	I	S	L	Y	G	W	T	R	N	P	L	I	E	Y	I	V	E	S	Y	G	S	Y	N		
B0FX61	V	V	T	Y	N	G	T	W	N	G	A	S	V	N	S	Y	L	S	V	Y	G	W	T	R	N	P	L	I	E	Y	I	V	V	E	S	F	G	T	Y	N	
Q0ZHI9	T	I	N	Y	G	G	S	F	S																																

4. Supplementary information

	β^{12}				$\alpha-1$				β^{13}				β^{14}																																						
	180		190		200		210		220		230		240																																						
C7YSL3*	..SS	GSVN	TGA	AHF	NA	WSNV	GLKL	.GSHD	YQILA	VEGYSSGS	ATMTV	S																																						
A2R5J7	..ST	GTVD	VS	AHF	QR	WKEL	GLTP	.GSPV	FQMV	TEGFSGQGY	LDFTV	SA																																						
Q2UFR7	..TT	GSVD	VTA	AHF	KRW	WKE	GLQP	.GNPV	FQMV	TEGFKSGSN	LDFQL	H																																						
Q2TYR4	..HT	GTVD	VAA	AHF	KRW	QEL	GLKP	.GNPV	FQMV	TEGFKSGSN	LDFTL	QK																																						
Q9AG99	.RTG	GTIT	AG	NHF	DA	WASK	GMNL	.GTHD	YQILA	TEGYSSGS	SNVTV	SE	GTISGGGNTGGG																																						
Q96TR7	..TS	GTVT	TGN	NHF	AY	WAKY	GFGN	.S.YN	FQVMA	VEAFSGTGS	ASVTI	VS																																						
Q12562	..TS	GTVT	TGN	NHF	AY	WAKY	GFGN	.S.YN	FQVMP	VEAFSGTGS	ASVTI	VS																																						
Q9UW17	..TS	GTVT	TGN	NHF	AY	WAKY	GFGN	.S.YN	FQVMA	VEAFSGTGS	ASVTI	VS																																						
A2I7V2	..TS	GTVT	VAN	NHF	EA	WAQH	GFGN	.SNFN	YQVVA	VEAWSGAGS	ASVTI	SS																																						
Q12549	..TS	GTVT	IAN	NHF	NF	WAQH	GFGN	.SNFN	YQVMA	VEAWN	GVGS	ASVTI	SS																																					
E3UN71	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SNFN	YQVMA	VEAWN	GAGS	ASVTI	ST																																					
A6N2L8	..TS	GTVT	IAN	NHF	NF	WAQH	GFGN	.SNFN	YQVMA	VEAWN	GAGS	ARVTI	SS																																					
Q45UD8	..TS	GTVT	IAN	NHF	NF	WAQH	GFGN	.SNFN	YQVMA	VEAWN	GAGS	ASVTI	SS																																					
Q2I0I8	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SDFN	YQVVA	VEAWSGAGS	ASVTI	SS																																						
Q12550	..TS	GTVT	VAI	NHF	NF	WAQH	GFGN	.SDFN	YQVMA	VEAWSGACS	ASVTI	SS																																						
P55331	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SDFN	YQVVA	VEAWSGAGS	ASVTI	SS																																						
P55328	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SDFN	YQVMA	VEAWSGAGS	ASVTI	SS																																						
P33557	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SDFN	YQVMA	VEAWSGAGS	ASVTI	SS																																						
P55329	..TS	GTVT	VAN	NHF	NF	WAQH	GFGN	.SDFN	YQVMA	VEAWSGAGS	ASVTI	SS																																						
Q6U894	G.NN	VSVT	FS	NHV	NA	WRNA	GMNL	GSSWS	YQVLA	TEGYSSGR	SNVTV	W																																						
Q71S35	G.NN	VSIT	FS	NHV	NA	WRNA	GMNL	GSSWS	YQVLA	TEGYSSGR	SNVTV	W																																						
Q43993	G.VN	STIT	FS	NHV	NA	WPSK	GMYL	GNSWS	YQVMA	TEGYSSGN	ANVTV	W																																						
Q1XGE6	G.VN	STIT	FS	NHV	NA	WASK	GMYL	GNSWS	YQVMA	TEGYSSGS	SNVTV	W																																						
Q6TLP3	G.TN	NITIT	FS	NHV	NA	WASK	GMNL	GSSWS	YQVLA	TEGYSSGY	SNVTV	W																																						
B5M610	G.IN	AAIT	FS	NHV	NA	WASK	GMNL	GNSWS	YQVLA	TEGYSSGS	SNVTV	W																																						
E0YL13	G.SN	AAIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVLA	TEGYKSSGS	SNVTV	W																																						
A2T401	G.SN	AKIT	FS	NHV	KA	WASK	GMNL	GSIWS	YQVLA	TEGYSSGS	SNVTV	W																																						
C5MTD6	G.SN	AKIT	FS	NHV	KA	WASK	GMNL	GSIWS	YQVLA	TEGYSSGS	SNVTV	W																																						
Q3HLJ4	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TEGYSSGS	SNVTV	W																																						
Q8RMN8	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TGGYSSGS	SNVTV	W																																						
Q45VU6	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TEGYSSGS	SNVTV	W																																						
Q45VU3	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TEGYSSGS	SNVTV	W																																						
P09850	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TEGYSSGS	SNVTV	W																																						
P18429	G.SN	ATIT	FS	NHV	NA	WASK	GMNL	GNSWA	YQVMA	TEGYSSGS	SNVTV	W																																						
Q9RMM4	..G.	GTIT	TGN	NHF	DA	WA	AH	GMP	L	GTFNY	.MILA	TEGYSSGS	SNITV	GD	SSGGDNG																																			
D7EZJ3	..G.	GTIT	TGN	NHF	DA	WA	RA	GM	Q	L	GSHDY	.MIMA	TEGYSSGS	SNITV	GG	TSG																																		
C6WIK2	..G.	GSIT	TGN	NHF	DA	W	SRY	GL	R	L	GDFRY	.YMIMA	TEGYSSGN	SNITV	S																																			
Q8GMV7	..S.	GTIT	TGN	NHF	DA	WA	RA	GM	N	L	GSHDY	.QIMA	TEGYSSGS	STVSI	SE	GGNPNP	GNPG																																		
Q9RI72	G.S.	GTIT	TGN	NHF	DA	WA	RA	GM	N	M	Q	F	RY	.YMIMA	TEGYSSGS	SNITV	SG																																	
Q56013	G.S.	GTIT	TGN	NHF	DA	WA	RA	GM	N	L	GQFQY	.YMIMA	TEGYSSGS	SNITV	SG																																			
Q59962	..G.	GSIT	TGN	NHF	DA	WA	RY	GMP	L	G	S	F	NY	.YMIMA	TEGYSSGS	SSISV	S																																	
P26515	..G.	GTIT	TGN	NHF	DA	WA	RA	GM	P	L	G	N	F	S	.YMIMA	TEGYSSGS	SSINV	GG	TGGG																															
B2LWN3	..VG	GTIT	TGN	NHF	DA	WA	S	V	GL	N	L	.G	T	H	N	.YQIMA	TEGYSSGS	SDITV	SE	GGSSSGGGG	..																														
B3PIN0	G.NI	ST	IT	F	A	N	H	A	N	F	W	A	S	K	GL	N	L	.GNHN	YQVLA	TEGYSSRGS	SDITV	S	QGGSSSGGNS	..																											
C5BU24	GNIS	GT	I	N	V	G	N	H	F	N	Y	W	A	S	K	GL	N	L	.GSHD	YMLA	TEGYSSGS	SDISV	SE	GGSSSSSSGS																											
Q92244	..AS	GTVT	FAN	HV	NA	WRNA	GLNL	GNQWN	YQILA	VEGYHSSGS	ASMTV	R																																						
Q43097	..TS	GTVQ	TG	CHF	DA	WA	RA	GL	N	V	NGDHY	.YQIVA	TEGYFSSGY	ARITV	AD	VG																																		
Q0ZBL0	..TS	GSVN	VGA	AHF	DA	WA	KS	GL	QL	.GRHN	.YQIVA	TEGYSSGS	SSITV	S																																				
Q00263	..SS	GSVN	MK	TH	DA	WA	AK	GM	KL	.G	T	H	N	.YQIVA	TEGYFSSGS	AQITV	N	CA																																
Q74716	..VG	GTVS	TG	VHF	NA	WR	SL	GM	PL	.G	T	Y	D	.YMIVA	TEGFRSSGS	ASITV	S																																	
Q9HFH0	..VG	GTVT	TAN	NHF	AA	W	K	A	L	GL	E	M	.G	T	Y	N	.YMIVS	TEGYESSGS	SSITV	S																														
A2QBA9	..IG	GTVT	TQN	NHF	DA	W	E	N	L	GL	E	L	.G	T	F	N	.YMIVA	TEGYESSGS	ATITV	E	TSSDSTEES																														
E0X4B3	..QS	GTVT	TK	NHF	DA	W	R	N	A	GL	QL	.G	N	F	D	.YMIVA	TEGYSSGS	ATITV	S																															
P87037	..VG	GTIT	AQ	NHF	DA	W	A	N	V	GL	QL	.G	T	H	N	.YMILA	TEGYKSSGS	ATITV	E																															
Q0Q592	..VG	GTIT	VK	NHF	DA	W	A	K	Y	G	Q	K	.G	T	H	N	.YQILA	TEGYSSGS	ASITV	E	GP																													
B0FX61	..VG	GSVT	MQ	NHF	DA	W	E	S	A	GL	KL	.G	T	H	N	.YQILA	TEGYSSGS	AKITV	E	QS																														
Q0ZHI9	..VG	GAVT	MQ	NHF	NA	W	A	Q	F	GM	N	L	.G	A	H	Y	.YQIVA	TEGYSSGP	SDIYV	Q	TQCKSLCDRGRV																														
Q0ZBK9	..SS	GAVT	F	G	N	H	V	Y	A	W	S	Q	K	GL	N	L	.G	T	H	N	.YQIMA	TEGYSSGS	SSITV	W																										
D3KT79	..VG	GTVT	TAN	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	ASITV	A																														
P48824	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQILA	TEGYSSGS	SSITV	Q																														
Q6QA21	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
F5CI28	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
Q3S401	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
A2Q7I0	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
P55330	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
Q2PU02	..VG	GTVT	TS	NHF	NA	W	A	K	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Q																														
Q9HFA4	..VG	GTVT	TGN	NHF	NA	W	A	K	Y	GL	T	L	.G	T	H	N	.YQIVA	TEGYSSGS	SAITV	Y																														
P55333	..TS	GSVT	TQN	NHF	DA	W	A	S	Q	L	GM	T	L	.G	T	H	N	.YQIVA	VEGYSSGS	ASITV	S																													
P55332	..TG	GTVT	TAN	NHF	NA	W	A	A	L	GM	R	L	.G	T	H	N	.YQIVA	TEGYSSGS	ASITV	Y																														
A1CCU0	..TG	GTVT	TAN	NHF	NA	W	A	A	L	GM	S	R	L	.G	T	H	N	.YQIVA	TEGYSSGS	SSITV	Y																													
Q4WGL1	..TG	GTVT	MAN	NHF	NA	W	S	R	L	GM	N	L	.G	T	H	N	.YQIVA	TEGYSSGS	ASITV	Y																														
Q69IG4	..TS	GHTI	VS	D	H	F	K	W	A	K	H	GW	I	G	.N	L	Y	E	V	A	L	N	A	E	G	W	S	S	G	I	A	D	V	T	K	L	D	V	Y	T	T	Q	K	S	N	P	T				
Q6QHA0	..QS	GTIS	I	T	E	H	F	K	Q	W	E	S	M	GL	K	L	G	.N	M	Y	E	A	K	F	L	V	E	A	G	G	T	G	W	L	E	F	T	Y	L	K	L	Q	E	E	K	R	N	..			
D4PAK8	..KS	GTIS	I	T	E	H	F	K	W	E	K	L	GM	K	L	G	D	N	M	Y	E	C	K	F	L	I	E	A	G	A	G	E	G	F	D	A	R	L	I	Q	F	Y	R	A	D	N	E	G	N	I	L
A4UWQ6	..QS	GTIS	VS	Q	H	F	D	W	A	S	L	GM	R	L	G	.S	L	Y	E	A	K	I	L	V	E	A	G	G	S	G	I	N	S	S	A	S	I													
E2JFA2	TKRE	GTIS	A	S	K	H	F	DA	W	K	K	C	L	E	L	G	.S	M	Y	E	V	A	L	T	I	E	G	Y	Q	S	N	G	S	A	V	I	Y	K	N	D	L	K	I	E	G	T	Y	S	E	A	T
P77853	..TS	GTVT	VTD	H	F	RA	W	A	N	R	GL	N	L	.G	.T	I	D	.Q	I	T	L	C	V	E	G	Y	Q	S	S	G	S	A	N	I	T	Q	N	I	F	S	Q	G	S	S	S	G	S	S	G		
P71737	..TS	GTIS	V	S	K	H	F	AA	W	E	S	K	GM	P	L	G	.K	M	H	E	T	A	F	N	I	E	G	Y	Q	S	S	G	K	A	D	V	N	S	M	S	I	N	I	G						
Q7SID8	..TS	GTVS	V	S	E	H	F	KK	W	E	S	L	GM	P	M	G	.K	M	Y	E	T	A	L	T	V	E	G	Y	Q	S	N	G	S	A	N	V	T	A	N	V	L	T	I	G	G	K	P	L			
E2IHA1	..TS	GTVS	V	S	A	H	F	KK	W	E	S	L	GM	P	M	G	.K	M	Y	E	T	A	F	T	V	E	G	Y	Q	S	S	G	S	A	N	V	M	T	N	Q	L	F	I	G	N					
Q8RMN7	..TS	GTVS	V	S	A	H	F	KK	W	E	S	L	GM	P	M	G	.K	M	Y	E	T	A	F	T	V	E	G	Y	Q	S	S	G	S	A	N	V	M	T	N	Q	L	F	I	G	N					
B5MOC6	..TS	GTVS	V	S	A	H	F	KK	W	E	S	L	GM	P	M	I	G	.K	M	Y	E	T	A	F	T	V	E	G	Y	Q	S	S	G	S	A	N	V	M	T	N	Q	L	F	I	G	N				

4. Supplementary information

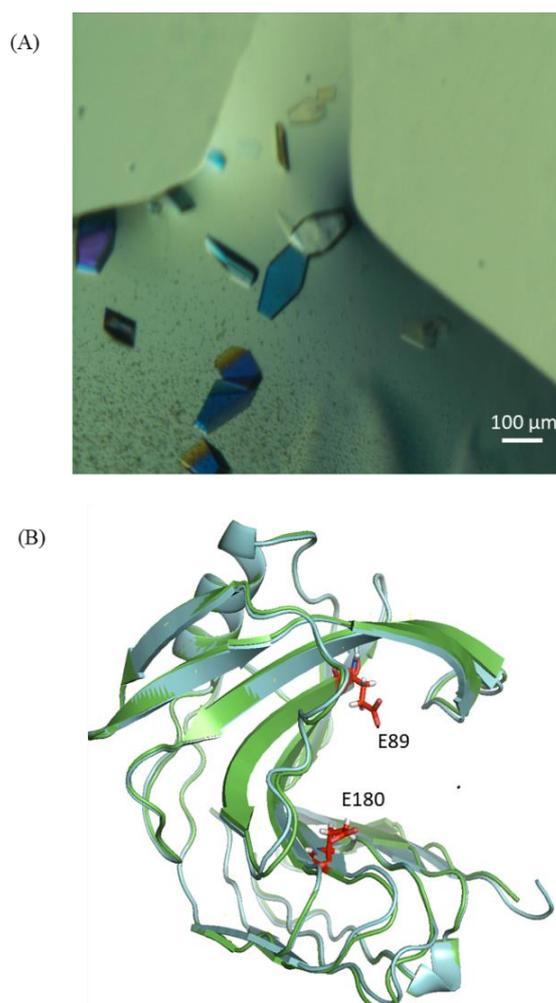
C7YSL3*
A2R5J7
Q2UFR7
Q2TYR4 GDTGGGNTGGGGT SNGCTVTATRSESWSDRFNVNYT
Q9AG99
Q96TR7
Q12562
Q9UW17
A2I7V2
Q12549
E3UN71
A6N2L8
Q45UD8
Q2I0I8
Q12550
P55331
P55328
P33557
P55329
Q6U894
Q71S35
Q43993
Q1XGE6
Q6TLP3
B5M6I0
E0YL13
A2T401
C5MTD6
Q3HLJ4
Q8RMN8
Q45VU6
Q45VU3
P09850
P18429 G. . GGGGGGGGGNTGGCTATLSAGEQWSDRYNLNVS
Q9RMM4 GGGGGGGGGGCTATLSAGERWDDRYNLNVS
D7EZZ3
C6WIK2 N. . PGNP GNP GNP GGGCVATLSAGQQWSDRYNLNVS
Q8GMV7
Q9RI72
Q56013
Q59962 . . . DSGGGDNGGGGGGCTATVSAGQKWGDRYNLDVS
P26515 . . . STSSSSGGGGTKSFTVRARGTAGGESITLRVN
B2LWN3 . . . SSSSSASGGGSKIIVVRARGTAGGESITLRVG
B3FIN0 SSSGGSSSSSSGGSGGTITVRARGNSGGENIALRVG
C5BU24
Q92244
O43097
Q0ZBL0
Q00263
O74716
Q9HFH0
A2QBA9
E0X4B3
P87037
Q0Q592
B0FX61
Q0ZHI9 TWRDVVC
Q0ZBK9
D3KT79
P48824
Q6QA21
F5CI28
Q3S401
A2Q7I0
P55330
Q2PU02
Q9HFA4
P55333
P55332
A1CCU0
Q4WG11
Q69IG4 TAARTTRTTARTTA. . RTTTRTKTLPNNKCSSKITAQGYKCCSNPNCE. IVYTDGDDGTW
Q6QHA0
D4PAK8 QITP
A4UWQ6
E2JFA2 DIE. VNVN. E
P77853 GSSGSTITTTRIECE. NMSLSGPYVSRIT. NPFNGIALYANGDTAR
P17137
Q7SID8
E2IHA1
Q8RMN7
B5MOC6



S. 4.3 Phylogenetic tree indicating the evolutionary relationships among GH11 xylanases. FASTA sequences of Fungal GH11 xylanases from sequence 1-46 indicated with red color and Bacterial GH11 xylanases from sequence 47-82 indicated with black color. NhGH11 is indicated with green color. The pattern of branching indicated the evolution of NhGH11 from common fungal ancestors. Branch points are indicated with blue dots and branches are labeled with a confidence score. The sequences with smaller confidence scores mean they evolve from a common ancestor.

S. 4.4. Active site mutant for soaking of substrates

The active site residues E89 and E180 of NhGH11 were mutated to alanines and protein was expressed and purified using procedures described for NhGH11 (section 2.4-2.7). The activity assay of active site mutant was performed by DNS assay¹⁵⁸ using the standard procedure described in section 2.15.1. The activity results showed that the active site mutant was 50% less active than NhGH11. Crystals of active site mutant could grow only in precipitant condition 1 (1 M ammonium sulfate, 100 mM sodium citrate pH 5.5), following the protocol described in section 2.13. Co-crystals of active site mutant with saccharidic substrates could not be obtained.



S. 4.4. (A) The active site mutant was crystallized by vapor diffusion-hanging drop by mixing equal volumes of 10 mg/mL protein and a precipitant solution containing 1 M ammonium sulfate, 100 mM sodium citrate pH 5.5. The streak seeding technique was applied utilizing the same seed stock that was applied for NhGH11 crystals. Mutant crystals were obtained at 20 °C after two days. (B) Superimposition of NhGH11 (green) and active site mutant (cyan). Mutated residues are shown with red color.

Table 4.1 Data-collection and refinement statistics for active site mutant

Data collection	
X-ray source	P11 beamline PETRA III, DESY
Detector	Pilatus
Space group	C 2 (No. 5)
Cell dimensions	
a, b, c (Å)	81.6, 38.4, 51.1
α, β, γ (°)	90.0, 91.1, 90.0
Wavelength (Å)	0.97
Resolution range (Å)	40.8-1.3 (1.28-1.3)
Total number of reflections	32537
Redundancy	3.1 (2.6)
Wilson B-factor (Å²)	8.59
R_{meas}	0.053 (0.49)
R_{merge}	0.045 (0.47)
CC_{1/2}	0.999 (0.857)
I/σI	15.57 (2.33)
Completeness (%)	85.12 (80.74)
Refinement	
Reflections used	82543 (6956)
Reflection used for R_{free}	1970 (174)
R_{work}	0.20
R_{free}	0.24
No. atoms	1863
Protein	1580
Water	278
Ligands	Not observed
Average B-factor (Å²)	20.78
R.m.s deviations	
Bond lengths (Å)	0.0074
Bond angles (°)	1.1
Ramachandran	
Favored (%)	95
Allowed (%)	5
Outliers (%)	0.00

5. Bibliography

1. Singla, A. *et al.* Bioethanol production from xylose: Problems and possibilities. *J. Biofuels* (2012), 3(1), 39-49.
2. Sánchez, C. Lignocellulosic residues: Biodegradation and bioconversion by fungi. *Biotechnol. Adv.* (2009), 27(2), 185-194.
3. Bhardwaj, N., Kumar, B., Agarwal, K., Chaturvedi, V. & Verma, P. Purification and characterization of a thermo-acid/alkali stable xylanases from *Aspergillus oryzae* LC1 and its application in Xylo-oligosaccharides production from lignocellulosic agricultural wastes. *Int. J. Biol. Macromol.* (2019), 122, 1191-1202.
4. Collins, T., Gerday, C. & Feller, G. Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiology Reviews* (2005), 29(1), 3-23.
5. Pathania, S., Sharma, N. & Handa, S. Optimization of culture conditions using response surface methodology for synergism production of cellulase , xylanase and pectinase by *Rhizopus delemar* F2 under solid state fermentation. *J. Pharmacogn. Phytochem.* (2017), 6(6):1872-1878.
6. Akhtar, N., Gupta, K., Goyal, D. & Goyal, A. Recent advances in pretreatment technologies for efficient hydrolysis of lignocellulosic biomass. *Environ. Prog. Sustain. Energy* (2016), 35(2), 489-511.
7. Sherwood, J. The significance of biomass in a circular economy. *Bioresour. Technol.* (2020), 300, 122755.
8. Loqué, D., Scheller, H. V. & Pauly, M. Engineering of plant cell walls for enhanced biofuel production. *Curr. Opin. Plant Biol.* (2015), 25, 151-161.
9. Van Den Berg, N. L. & Ashmore, M. Introduction Nitrogen Pollution in a Global Perspective Nitrogen Ecotoxicity in Terrestrial Ecosystems. *Encycl. Ecol.* (2008).
10. Axelsson, L. *et al.* Perspective: Jatropha cultivation in southern India: Assessing farmers' experiences. *Biofuels, Bioprod. Biorefining* (2012), 6, 246-256.
11. Vasić, K., Knez, Ž. & Leitgeb, M. Bioethanol Production by Enzymatic Hydrolysis from Different Lignocellulosic Sources. *Molecules* (2021), 26(3), 753.
12. Sorek, N., Yeats, T. H., Szemenyei, H., Youngs, H. & Somerville, C. R. The

- implications of lignocellulosic biomass chemical composition for the production of advanced biofuels. *Bioscience* (2014), 64, 192-201.
13. Aro, E. M. From first generation biofuels to advanced solar biofuels. *Ambio* (2016), 45(1), 24-31.
 14. Ananthi, V. *et al.* A realistic scenario on microalgae based biodiesel production: Third generation biofuel. *Fuel* (2021), 284, 118965.
 15. Ale, S., Femeena, P. V., Mehan, S. & Cibin, R. Environmental impacts of bioenergy crop production and benefits of multifunctional bioenergy systems. *Bioenergy with Carbon Capture Storage Using Nat. Resour. Sustain. Academic Press.* (2019), 195-217.
 16. Cameron, D. E., Bashor, C. J. & Collins, J. J. A brief history of synthetic biology. *Nature Reviews, Microbiology* (2014), 12(5), 381-390.
 17. Bessou, C., Ferchaud, F., Gabrielle, B. & Mary, B. Biofuels, greenhouse gases and climate change. *Sustainable Agriculture* (2009), 2, 365-468.
 18. Bairamzadeh, S., Saidi-Mehrabad, M. & Pishvaei, M. S. Modelling different types of uncertainty in biofuel supply network design and planning: A robust optimization approach. *Renew. Energy* (2018), 116, 500-517.
 19. Ahorsu, R., Medina, F. & Constantí, M. Significance and challenges of biomass as a suitable feedstock for bioenergy and biochemical production: A review. *Energies* (2018), 11(12), 3366.
 20. Hawthorne, B. T., Rees-George, J. & Broadhurst, P. G. Mating behaviour and pathogenicity of new zealand isolates of *Nectria haematococca* (*Fusarium solani*). *New Zeal. J. Crop Hortic. Sci.* (1992), 20(1), 51-57.
 21. Hawthorne, B. T., Ball, R. D. & Rees-George, J. Genetic analysis of variation of pathogenicity in *Nectria haematococca* (*Fusarium solani*) on Cucurbita sp. *Mycol. Res.* (1994), 20(1), 51-57.
 22. Anaissie, E. & Nucci, M. Cutaneous infection by *Fusarium* species in healthy and immunocompromised hosts: Implications for diagnosis and management. *Clin. Infect. Dis.* (2002), 35, 909–920.
 23. Zhang, N. *et al.* Members of the *Fusarium solani* species complex that cause

- infections in both humans and plants are common in the environment. *J. Clin. Microbiol.* (2006), 44, 2185-2190.
24. Chang, D. C. *et al.* Multistate outbreak of *Fusarium* keratitis associated with use of a contact lens solution. *J. Am. Med. Assoc.* (2006), 296(8), 953-963.
25. Coleman, J. J. *et al.* The genome of *Nectria haematococca*: Contribution of supernumerary chromosomes to gene expansion. *PLoS Genet.* (2009), 5(8), e1000618.
26. Chen, S. Y., Dickson, D. W. & Mitchell, D. J. Pathogenicity of fungi to eggs of *Heterodera glycines*. *J. Nematol.* (1996), 28(2), 148.
27. Chettri, D., Verma, A. K. A. K. & Verma, A. K. A. K. Innovations in CAZyme gene diversity and its modification for biorefinery applications. *Biotechnol. Reports* 28, (2020), e00525.
28. Paës, G., Berrin, J. G. & Beaugrand, J. GH11 xylanases: Structure/ function/ properties relationships and applications. *Biotechnol. Adv.* (2012), 30, 564–592.
29. Luo, H. *et al.* Cloning, expression and characterization of a novel acidic xylanase, XYL11B, from the acidophilic fungus *Bispora* sp. MEY-1. *Enzyme Microb. Technol.* (2009), 45, 126-133.
30. Hayashi, H. *et al.* Nucleotide sequences of two contiguous and highly homologous xylanase genes *xynA* and *xynB* and characterization of *XynA* from *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* (1999), 51(3), 348-357.
31. Krishnamurthy, S. & Vithayathil, P. J. Purification and characterization of endo-1,4- β -xylanase from *Paecilomyces varioti* Bainier. *J. Ferment. Bioeng.* (1989), 67, 77-82.
32. Li, N. *et al.* Cloning, expression, and characterization of a new *Streptomyces* sp. S27 xylanase for which xylobiose is the main hydrolysis product. *Appl. Biochem. Biotechnol.* (2009), 159(2), 521-531.
33. Gomes, J. *et al.* Production of high level of cellulase-free and thermostable xylanase by a wild strain of *Thermomyces lanuginosus* using beechwood xylan. *J. Biotechnol.* (1993) 30, 283-297.
34. Shrinivas, D., Savitha, G., Raviranjana, K. & Naik, G. R. A highly thermostable

- alkaline cellulase-free xylanase from thermoalkalophilic *Bacillus* sp. JB 99 suitable for paper and pulp industry: Purification and characterization. *Appl. Biochem. Biotechnol.* (2010), 162(7), 2049-2057.
35. Morris, D. D., Gibbs, M. D., Ford, M., Thomas, J. & Bergquist, P. L. Family 10 and 11 xylanase genes from *Caldicellulosiruptor* sp. strain Rt69B.1. *Extremophiles* (1999), 3, 103-111.
36. Sakka, K. *et al.* Purification and Characterization of Xylanase A from *Clostridium stercorarium* F-9 and a Recombinant Escherichia coli. *Biosci. Biotechnol. Biochem.* (1994), 58, 1496-1499.
37. Monti, R., Terenzi, H. F. & Jorge, J. A. Purification and properties of an extracellular xylanase from the thermophilic fungus *Humicola grisea* var. *thermoidea*. *Can. J. Microbiol.* (1991), 37(9), 675-681.
38. Lee, S. F. & Forsberg, C. W. Purification and characterization of an α -L-arabinofuranosidase from *Clostridium acetobutylicum* ATCC 824. *Can. J. Microbiol.* (1991), 57(1) 212-218.
39. Cheng, Y. F., Yang, C. H. & Liu, W. H. Cloning and expression of Thermobifida xylanase gene in the methylotrophic yeast *Pichia pastoris*. *Enzyme Microb. Technol.* (2005) 37(5), 541-546.
40. Hakulinen, N., Turunen, O., Jänis, J., Leisola, M. & Rouvinen, J. Three-dimensional structures of thermophilic β -1,4-xylanases from *Chaetomium thermophilum* and *Nonomuraea flexuosa*: Comparison of twelve xylanases in relation to their thermal stability. *Eur. J. Biochem.* (2003), 270, 1399-1412.
41. Morris, D. D. *et al.* Cloning of the xynB gene from *Dictyoglomus thermophilum* Rt46b.1 and action of the gene product on kraft pulp. *Appl. Environ. Microbiol.* (1998), 64(5), 1759-1765.
42. Long, L. *et al.* Thermostable xylanase-aided two-stage hydrolysis approach enhances sugar release of pretreated lignocellulosic biomass. *Bioresour. Technol.* (2018), 64(5), 334-338.
43. Polizeli, M. L. T. M. *et al.* Xylanases from fungi: Properties and industrial applications. *Applied Microbiology and Biotechnology* (2005), 67(5), 577-591.
44. Saha, B. C. Production, purification and properties of xylanase from a newly
-

- isolated *Fusarium proliferatum*. *Process Biochem.* (2002), 37 (11) 1279-1284.
45. Dwivedi, P., Vivekanand, V., Ganguly, R. & Singh, R. P. Parthenium sp. as a plant biomass for the production of alkalitolerant xylanase from mutant *Penicillium oxalicum* SAU_E-3.510 in submerged fermentation. *Biomass and Bioenergy* (2009), 33, 581-588
 46. Pal, A. & Khanum, F. Production and extraction optimization of xylanase from *Aspergillus niger* DFR-5 through solid-state-fermentation. *Bioresour. Technol.* (2010), 101, 7563–7569
 47. Zouari Ayadi, D. *et al.* Improvement of *Trichoderma reesei* xylanase II thermal stability by serine to threonine surface mutations. *Int. J. Biol. Macromol.* (2015), 72, 163-170.
 48. Yang, S. Q. *et al.* High-level of xylanase production by the thermophilic *Paecilomyces thermophila* J18 on wheat straw in solid-state fermentation. *Bioresour. Technol.* (2006), 97, 1794-1800.
 49. Du, Y. *et al.* Characterization of three novel thermophilic xylanases from *Humicola insolens* Y1 with application potentials in the brewing industry. *Bioresour. Technol.* (2013), 130, 161-167.
 50. Chutani, P. & Sharma, K. K. Biochemical evaluation of xylanases from various filamentous fungi and their application for the deinking of ozone treated newspaper pulp. *Carbohydr. Polym.* (2015), 127, 54-63.
 51. Ding, C., Li, M. & Hu, Y. High-activity production of xylanase by *Pichia stipitis*: Purification, characterization, kinetic evaluation and xylooligosaccharides production. *International Journal of Biological Macromolecules* (2018), Elsevier vol. 117.
 52. Sadaf, A. & Khare, S. K. Production of *Sporotrichum thermophile* xylanase by solid state fermentation utilizing deoiled *Jatropha curcas* seed cake and its application in xylooligosachharide synthesis. *Bioresour. Technol.* (2014), 153, 126-130.
 53. Chang, A. *et al.* BRENDA, the ELIXIR core data resource in 2021: New developments and updates. *Nucleic Acids Res.* (2021), Vol. 49, Database issue.
 54. Goesaert, H. *et al.* Wheat flour constituents: How they impact bread quality, and how to impact their functionality. *Trends in Food Science and Technology* (2005),
-

- 16, 12-30.
55. Verjans, P., Dornez, E., Delcour, J. A. & Courtin, C. M. Selectivity for water-unextractable arabinoxylan and inhibition sensitivity govern the strong bread improving potential of an acidophilic GH11 *Aureobasidium pullulans* xylanase. *Food Chem.* (2010), 123, 331-337
56. Kumar, D., Kumar, S. S., Kumar, J., Kumar, O., Mishra, S. V., Kumar, R., & Malyan, S. K. Xylanases and their industrial applications: a review. *Biochem Cell Arch.* (2017) 17(1), 353-360.
57. Awalgaonkar, G., Sarkar, S., Bankar, S. & Singhal, R. S. Xylanase as a processing aid for papads, an Indian traditional food based on black gram. *LWT - Food Sci. Technol.* (2015), 62, 1148-1153.
58. Frederix, S. A., Courtin, C. M. & Delcour, J. A. Substrate selectivity and inhibitor sensitivity affect xylanase functionality in wheat flour gluten-starch separation. *J. Cereal Sci.* (2004), 40, 41-49.
59. Van Der Borght, A., Goesaert, H., Veraverbeke, W. S. & Delcour, J. A. Fractionation of wheat and wheat flour into starch and gluten: Overview of the main processes and the factors involved. *J. Cereal Sci.* (2005), 41, 221-237.
60. Gao, F., Jiang, Y., Zhou, G. H. & Han, Z. K. The effects of xylanase supplementation on growth, digestion, circulating hormone and metabolite levels, immunity and gut microflora in cockerels fed on wheat-based diets. *Br. Poult. Sci.* (2007), 48(4), 480-488.
61. Vandeplass, S., Dauphin, R. D., Thonart, P., Théwis, A. & Beckers, Y. Effect of the bacterial or fungal origin of exogenous xylanases supplemented to a wheat-based diet on performance of broiler chickens and nutrient digestibility of the diet. *Can. J. Anim. Sci.* (2010), 90(2), 221-228.
62. Van Craeyveld, V. *et al.* Wheat bran AX properties and choice of xylanase affect enzymic production of wheat bran-derived arabinoxylan-oligosaccharides. *Cereal Chem.* (2010), 87(4) 283-291.
63. Mussatto, S. I. & Mancilha, I. M. Non-digestible oligosaccharides: A review. *Carbohydrate Polymers* (2007), 8, 587-597.
64. Samanta, A. K. *et al.* Xylooligosaccharides as prebiotics from agricultural by-
-

- products: Production and applications. *Bioactive Carbohydrates and Dietary Fibre* (2015), 5(1), 62-71.
65. Van Craeyveld, V. *et al.* Structurally different wheat-derived arabinoxylooligosaccharides have different prebiotic and fermentation properties in rats. *J. Nutr.* (2008), 138(12), 2348-2355.
66. Walia, A., Guleria, S., Mehta, P., Chauhan, A. & Parkash, J. Microbial xylanases and their industrial application in pulp and paper biobleaching: a review. *3 Biotech* (2017), 7(1), 11.
67. Sharma, D., Chaudhary, R., Kaur, J. & Arya, S. K. Greener approach for pulp and paper industry by Xylanase and Laccase. *Biocatalysis and Agricultural Biotechnology* (2020), 25, 101604.
68. Kenealy, W. R. & Jeffries, T. W. Enzyme processes for pulp and paper: A review of recent developments. In *Wood deterioration and preservation: advances in our changing world*. Washington, DC: American Chemical Society: Distributed by Oxford University Press, in *ACS Symposium Series* (2003), 845, 210-239.
69. Ibarra, D., Köpcke, V., Larsson, P. T., Jääskeläinen, A. S. & Ek, M. Combination of alkaline and enzymatic treatments as a process for upgrading sisal paper-grade pulp to dissolving-grade pulp. *Bioresour. Technol.* (2010), 101, 7416-7423.
70. Olsen, H. S. & Falholt, P. The role of enzymes in modern detergency. *J. Surfactants Deterg.* (1998), 1(4), 555-567.
71. Das, T. & Ramaswamy, G. N. Enzyme treatment of wool and specialty hair fibers. *Textile Research Journal* (2006), 76, 126-133.
72. Espejo, F. Role of commercial enzymes in wine production: a critical review of recent research. *Journal of Food Science and Technology* (2021), 58(1), 9-21
73. Olfa, E., Mondher, M., Issam, S., Ferid, L. & Nejib, M. M. Induction, properties and application of xylanase activity from sclerotinia sclerotiorum S2 fungus. *J. Food Biochem.* (2007), 31, 96-107.
74. Strauss, M. L. A., Jolly, N. P., Lambrechts, M. G. & Van Rensburg, P. Screening for the production of extracellular hydrolytic enzymes by non-Saccharomyces wine yeasts. *J. Appl. Microbiol.* (2001), 91(1), 182-190.

75. Morgan Fröling 1,2 Andrew Peterson² and Jefferson W TesterMargeot, A., Hahn-Hagerdal, B., Edlund, M., Slade, R. & Monot, F. New improvements for lignocellulosic ethanol. *Current Opinion in Biotechnology* (2009), 20(3), 372-380.
76. Mosier, N. *et al.* Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* (2005), 96(6), 673-686.
77. Aguilar-Reynosa, A. *et al.* Microwave heating processing as alternative of pretreatment in second-generation biorefinery: An overview. *Energy Conversion and Management* (2017), 136, 50-65.
78. Fröling, M., Peterson, A., Jefferson, W. *Hydrothermal Processing in Biorefineries* (2017).
79. Badiei, M., Asim, N., Jahim, J. M. & Sopian, K. Comparison of Chemical Pretreatment Methods for Cellulosic Biomass. *APCBEE Procedia* (2014), 9, 170-174.
80. Bhatia, S. K., Kim, S. H., Yoon, J. J. & Yang, Y. H. Current status and strategies for second generation biofuel production using microbial systems. *Energy Convers. Manag.* (2017), 148, 1142-1156.
81. Wan, C. & Li, Y. Fungal pretreatment of lignocellulosic biomass. *Biotechnology Advances* (2012), 30(6), 1447-1457.
82. Saini, J. K., Saini, R. & Tewari, L. Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech* (2015), 5, 337-353.
83. Rocha-Meneses, L., Raud, M., Orupöld, K. & Kikas, T. Second-generation bioethanol production: A review of strategies for waste valorisation. *Agron. Res.* (2017), 15, 830-847.
84. Eijsink, V. G. H. *et al.* On the functional characterization of lytic polysaccharide monoxygenases (LPMOs). *Biotechnology for Biofuels* (2019), 12(1), 1-16.
85. Soares Júnior, F. L. *et al.* Endo-and exoglucanase activities in bacteria from mangrove sediment. *Brazilian J. Microbiol.* (2013), 44, 969-976.
86. Edwards, S., Chaplin, M. F., Blackwood, A. D. & Dettmar, P. W. Primary structure of arabinoxylans of ispaghula husk and wheat bran. *Proc. Nutr. Soc.* (2003), 62,

- 217-222.
87. Gómez-García, R. *et al.* Production of a xylanase by *Trichoderma harzianum* (Hypocrea lixii) in solid-state fermentation and its recovery by an aqueous two-phase system. *Can. J. Biotechnol.* (2018), 2(2), 108-115.
 88. Romero-Fernández, M. *et al.* Designing continuous flow reaction of xylan hydrolysis for xylooligosaccharides production in packed-bed reactors using xylanase immobilized on methacrylic polymer-based supports. *Bioresour. Technol.* (2018), 266, 249-258.
 89. Foreman, P. K. *et al.* Transcriptional regulation of biomass-degrading enzymes in the filamentous fungus *Trichoderma reesei*. *J. Biol. Chem.* (2003), 278(34), 31988-31997.
 90. Sarkar, N., Ghosh, S. K., Bannerjee, S. & Aikat, K. Bioethanol production from agricultural wastes: An overview. *Renewable Energy* (2012), 37(1), 19-27.
 91. Van Maris, A. J. A. *et al.* Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology* (2006), 90(4), 391-418.
 92. Matsushika, A., Inoue, H., Kodaki, T. & Sawayama, S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: Current state and perspectives. *Applied Microbiology and Biotechnology* (2009) 84 (1) 37-53.
 93. Delgenes, J. P., Moletta, R. & Navarro, J. M. The effect of aeration on D-xylose fermentation by *Pachysolen tannophilus*, *Pichia stipitis*, *Kluyveromyces marxianus* and *Candida shehatae*. *Biotechnol. Lett.* (1986), 8, 897-900.
 94. Nair, R. B., Lundin, M., Brandberg, T., Lennartsson, P. R. & Taherzadeh, M. J. Dilute phosphoric acid pretreatment of wheat bran for enzymatic hydrolysis and subsequent ethanol production by edible fungi *Neurospora intermedia*. *Ind. Crops Prod.* (2015), 69, 314-323.
 95. Echegaray, O. F. *et al.* Fed-batch culture of *Saccharomyces cerevisiae* in sugar-cane blackstrap molasses: Invertase activity of intact cells in ethanol fermentation. *Biomass and Bioenergy* (2000), 19(1), 39-50.
 96. Gírio, F. M. *et al.* Hemicelluloses for fuel ethanol: A review. *Bioresource*
-

- Technology* (2010), 101, 4775-4800.
97. Lin, Y. & Tanaka, S. Ethanol fermentation from biomass resources: Current state and prospects. *Applied Microbiology and Biotechnology* (2006), 69(6), 627-642.
 98. Luedeking, R. & Piret, E. L. Kinetic study of the lactic acid fermentation. Batch process at controlled pH. *Biotechnol. Bioeng.* (2000), 67(6), 636-644.
 99. Brethauer, S. & Wyman, C. E. Review: Continuous hydrolysis and fermentation for cellulosic ethanol production. *Bioresour. Technol.* (2010), 101(13), 4862-4874.
 100. Saarela, U., Leiviskä, K. & Juuso, E. Modelling of a fed-batch fermentation process. *Most* (2003), 1-21.
 101. Amornraksa, S., Subsaipin, I., Simasatitkul, L. & Assabumrungrat, S. Systematic design of separation process for bioethanol production from corn stover. *BMC Chem. Eng.* (2020), 2(1), 1-16.
 102. Tgarguifa, A., Abderafi, S. & Bounahmidi, T. Modeling and optimization of distillation to produce bioethanol. *Energy Procedia* (2017), 139, 43-48.
 103. Segovia-Hernandez, J. G. Alternative Schemes for the Purification of Bioethanol: A Comparative Study. *Recent Advances. Petrochemical Science.* (2018), 4(2) Rapsci.Ms.Id.555631.
 104. Röntgen, W. C. Ueber eine neue Art von Strahlen. *Ann. Phys.* (1898).
 105. Chen, H., Rogalski, M. M. & Anker, J. N. Advances in functional X-ray imaging techniques and contrast agents. *Physical Chemistry Chemical Physics* (2012), 14(39), 13469-13486.
 106. Friedrich, W., Knipping, P. & Laue, M. Interferenzerscheinungen bei Röntgenstrahlen. *Ann. Phys.* (1913), 346 (10), 971-988.
 107. Ewald, P. P. & Barlow, M. Fifty Years of X-Ray Diffraction. Fifty Years of X-Ray Diffraction: Dedicated to the International Union of Crystallography on the Occasion of the Commemoration Meeting in Munich July 1962. Springer Science & Business Media. *Phys. Today* (1963).
 108. Kendrew, J. C. *et al.* A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* (1958), 181, 662-666.
 109. Perutz, M. F. *et al.* Structure of Hæmoglobin: A three-dimensional fourier synthesis
-

- at 5.5-Å resolution, obtained by X-ray analysis. *Nature* (1960), 185, 416-422.
110. Nakagawa, A. Synchrotron Radiation. *Yuki Gosei Kagaku Kyokaishi/Journal Synth. Org. Chem.* (1996) Nakagawa, A. (1996), 54(5), 384-394.
111. Schmidt, M. Mix and inject: Reaction initiation by diffusion for time-resolved macromolecular crystallography. *Adv. Condens. Matter Phys.* (2013), 10, 16.
112. Hendrickson, W. A. Synchrotron crystallography. *Trends in Biochemical Sciences* (2000).
113. Als-Nielsen, J. & McMorrow, D. Elements of modern X-ray physics. *Phys. Today* (2002), 55, 63.
114. Helliwell, J. R. & Mitchell, E. P. Synchrotron radiation macromolecular crystallography: Science and spin-offs. *IUCrJ* (2015), 2, 283-291.
115. Wang, H. W. & Wang, J. W. How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Science* (2017), 26, 32-39.
116. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. - Sect. D Biol. Crystallogr.* (2001), 57, 1373-1382.
117. Dauter, Z. SAD/MAD Phasing. in *Advancing Methods for Biomolecular Crystallography*, eds. Read, R., Urzhumtsev, A. G. & Lunin, V. Y., Springer Netherlands (2013), 135-149.
118. Evans, P. & McCoy, A. An introduction to molecular replacement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2007), 64, 1-10.
119. Chapman, H. N. *et al.* Femtosecond X-ray protein nanocrystallography. *Nature* (2011), 470(7332), 73-77.
120. Beyerlein, K. R. *et al.* Mix-and-diffuse serial synchrotron crystallography. *IUCrJ* (2017), 4, 769-777.
121. Boutet, S. *et al.* High-Resolution Protein Structure Determination by Serial Femtosecond Crystallography. *Science* (2012), 337(6092), 362-364.
122. Gati, C. *et al.* Atomic structure of granulins determined from native nanocrystalline granulovirus using an X-ray free-electron laser. *Proc. Natl. Acad. Sci. U. S. A.* (2017), 114(9), 2247-2252.
123. Whistler, R. L. & Masak, E. Enzymatic Hydrolysis of Xylan. *J. Am. Chem. Soc.*
-

- (1955), **77**, 1241-1243.
124. Derewenda, U. *et al.* Crystal structure, at 2.6-Å resolution, of the *Streptomyces lividans* xylanase A, a member of the F family of β -1,4-D-glycanases. *J. Biol. Chem.* (1994), 269, 20811-20814.
125. White, A., Withers, S. G., Gilkes, N. R. & Rose, D. R. Crystal Structure of the Catalytic Domain of the β -1,4-Glycanase Cex from *Cellulomonas fimi*. *Biochemistry* (1994), 33(42), 12546-12552.
126. Comtat, J. Isolation, properties, and postulated role of some of the xylanases from the basidiomycete *Sporotrichum dimorphosporum*. *Carbohydr. Res.* (1983), 118, 215-231.
127. Fernandes, A. C. *et al.* Homologous xylanases from *Clostridium thermocellum*: Evidence for bi-functional activity, synergism between xylanase catalytic modules and the presence of xylan-binding domains in enzyme complexes. *Biochem. J.* (1999), 342, 105-110.
128. Biely, P., Vrřanská, M., Tenkanen, M. & Kluepfel, D. Endo- β -1,4-xylanase families: Differences in catalytic properties. *J. Biotechnol.* (1997), 57, 151-166.
129. Li, K. *et al.* Relationships between activities of xylanases and xylan structures. *Enzyme Microb. Technol.* (2000), 27, 89-94.
130. Nurizzo, D. *et al.* *Cellvibrio japonicus* α -l-arabinanase 43a has a novel five-blade β -propeller fold. *Nat. Struct. Biol.* (2002), 9(9), 665-668.
131. Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evolutionary Biology* (2012), 12,186.
132. Collins, T. *et al.* A novel family 8 xylanase, functional and physicochemical characterization. *J. Biol. Chem.* (2002), 277, 35133-35139.
133. Kulkarni, N., Shendye, A. & Rao, M. Molecular and biotechnological aspects of xylanases. *FEMS Microbiology Reviews* (1999), 23(4), 411-456.
134. Charnock, S. J. *et al.* The topology of the substrate binding clefts of glycosyl hydrolase family 10 xylanases are not conserved. *J. Biol. Chem.* (1999), 273, 32187-32199

135. Wan, Q. *et al.* X-ray crystallographic studies of family 11 xylanase Michaelis and product complexes: Implications for the catalytic mechanism. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2014), 70, 11-23.
136. Tahir, T. A. *et al.* Specific characterization of substrate and inhibitor binding sites of a glycosyl hydrolase family 11 xylanase from *Aspergillus niger*. *J. Biol. Chem.* (2002), 277, 44035-44043.
137. Debyser, W., Derdelinckx, G. & Delcour, J. A. Arabinoxylan solubilization and inhibition of the barley malt xylanolytic system by wheat during mashing with wheat wholemeal adjunct: Evidence for a new class of enzyme inhibitors in wheat. *J. Am. Soc. Brew. Chem.* (1997), 55, 153-156.
138. Juge, N., Payan, F. & Williamson, G. XIP-I, a xylanase inhibitor protein from wheat: A novel protein function. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2004), 1696(2), 203-211.
139. Pollet, A. *et al.* Identification of structural determinants for inhibition strength and specificity of wheat xylanase inhibitors TAXI-IA and TAXI-IIA. *FEBS J.* (2009), 276(14), 3916-3927.
140. Gebruers, K. *et al.* *Triticum aestivum* L. endoxylanase inhibitor (TAXI) consists of two inhibitors, TAXI I and TAXI II, with different specificities. *Biochem. J.* (2001), 353(2), 239-244.
141. Gebruers, K. *et al.* Properties of TAXI-type endoxylanase inhibitors. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2004), 1696(2), 213-221.
142. Kabel, M. A., Bos, G., Zeevalking, J., Voragen, A. G. J. & Schols, H. A. Effect of pretreatment severity on xylan solubility and enzymatic breakdown of the remaining cellulose from wheat straw. *Bioresour. Technol.* (2007), 98(10), 2034-2042.
143. Ximenes, E., Kim, Y., Mosier, N., Dien, B. & Ladisch, M. Deactivation of cellulases by phenols. *Enzyme Microb. Technol.* (2011), 48, 54-60.
144. Palmqvist, E. & Hahn-Hägerdal, B. Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. *Bioresour. Technol.* (2000), 74, 25-33.
145. Jönsson, L. J., Alriksson, B. & Nilvebrant, N. O. Bioconversion of lignocellulose: Inhibitors and detoxification. *Biotechnol. Biofuels* (2013), 6, 1-10.

-
146. Kim, Y., Ximenes, E., Mosier, N. S. & Ladisch, M. R. Soluble inhibitors/deactivators of cellulase enzymes from lignocellulosic biomass. *Enzyme Microb. Technol.* (2011), 48, 408-415.
 147. Ximenes, E., Kim, Y., Mosier, N., Dien, B. & Ladisch, M. Inhibition of cellulases by phenols. *Enzyme Microb. Technol.* (2010), 46, 170-176.
 148. Morrison, D., van Dyk, J. S. & Pletschke, B. I. The effect of alcohols, lignin and phenolic compounds on the enzyme activity of *Clostridium cellulovorans* XynA. *BioResources* (2011), 6, 3132-3141.
 149. Pan, X. Role of functional groups in lignin inhibition of enzymatic hydrolysis of cellulose to glucose. *J. Biobased Mater. Bioenergy* (2008), 2, 25-32.
 150. Payan, F. *et al.* The dual nature of the wheat xylanase protein inhibitor XIP-I: Structural basis for the inhibition of family 10 and family 11 xylanases. *J. Biol. Chem.* (2004), 279, 36029-36037.
 151. Duong-Ly, K. C. & Gabelli, S. B. Affinity Purification of a Recombinant Protein Expressed as a Fusion with the Maltose-Binding Protein (MBP) Tag. *Methods in Enzymology* (2015), 559, 17-26.
 152. Brunelle, J. L. & Green, R. One-dimensional SDS-polyacrylamide gel electrophoresis (*1D SDS-PAGE*). *Methods in Enzymology* (2014) 541, 151-159.
 153. Kümmel, R. P. W. Atkins: Physikalische Chemie. Weinheim, VCH Verlagsgesellschaft, ISBN 3-527-25913-9. *Acta Hydrochim. Hydrobiol.* (1989), 2, 227-232.
 154. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta - Proteins and Proteomics* (2005), 1751(2), 119-139.
 155. Yang, J. T., Wu, C. S. C. & Martinez, H. M. Calculation of protein conformation from Circular Dichroism. *Methods Enzymol.* (1986), 130, 208-269.
 156. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* (2007), 1(6), 2876.
 157. Kabsch, W. *et al.* XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2010), 66(2), 125-132.
 158. Miller, G. L. Use of Dinitrosalicylic acid reagent for determination of reducing
-

- sugar. *Anal. Chem.* (1959), 31(3), 426-428.
159. Huynh, K. & Partch, C. L. Analysis of protein stability and ligand interactions by thermal shift assay. *Curr. Protoc. protein Sci.* (2015), 79(1), 28-9.
160. Palm, G. J. *et al.* The structural basis for spectral variations in green fluorescent protein. *Nat. Struct. Biol.* (1997), 4(5), 361-365.
161. Derewenda, Z. S. Rational protein crystallization by mutational surface engineering. *Structure* (2004), 12(4), 529-535.
162. Doye, J. P. K., Louis, A. A. & Vendruscolo, M. Inhibition of protein crystallization by evolutionary negative design. *Physical Biology* (2004), 1(1), 9.
163. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S. & Eisenberg, D. Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci.* (2007), 16(8), 1569-1576.
164. Cooper, D. R. *et al.* Protein crystallization by surface entropy reduction: Optimization of the SER strategy. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2007), 63, 636-645.
165. Loll, P. J., Xu, P., Schmidt, J. T. & Melideo, S. L. Enhancing ubiquitin crystallization through surface-entropy reduction. *Acta Crystallogr. Sect. FStructural Biol. Commun.* (2014), 70, 1434-1442.
166. Asherie, N. Protein crystallization and phase diagrams. *Methods* (2004), 34(3), 266-272.
167. Vekilov, P. G. Nucleation. *Cryst. Growth Des.* (2010), 34(3), 266-272.
168. Oswald, C., Smits, S. H. J., Bremer, E. & Schmitt, L. Microseeding - A powerful tool for crystallizing proteins complexed with hydrolyzable substrates. *Int. J. Mol. Sci.* (2008), 9(7), 1131-1141.
169. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2004), 60, 2126-2132.
170. Merritt, E. A. Expanding the model: Anisotropic displacement parameters in protein structure refinement. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (1999), 55, 1109-1117.
171. Wang, J., Dauter, M., Alkire, R., Joachimiak, A. & Dauter, Z. Triclinic lysozyme at
-

-
- 0.65 Å resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2007), 63, 1254-1268
172. Schmidt, A., Teeter, M., Weckert, E. & Lamzin, V. S. Crystal structure of small protein crambin at 0.48 Å resolution. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* (2011), 67, 424-428.
173. Laulumaa, S. & Kursula, P. Sub-atomic resolution crystal structures reveal conserved geometric outliers at functional sites. *Molecules* (2019), 24(17), 3044.
174. Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. The PDB-REDO server for macromolecular structure model optimization. *IUCrJ* (2014), 1(4), 213-220.
175. McCarthy, A. A., Morris, D. D., Bergquist, P. L. & Baker, E. N. Structure of XynB, a highly thermostable β -1,4-xylanase from *Dictyoglomus thermophilum* Rt46B.1, at 1.8 Å resolution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* (2000), 56, 1367-1375.
176. Moers, K. *et al.* Alteration of *Bacillus subtilis* XynA endoxylanase substrate selectivity by site-directed mutagenesis. *Enzyme Microb. Technol.* (2007), 41, 85-91.
177. Taroni, C., Jones, S. & Thornton, J. M. Analysis and prediction of carbohydrate binding sites. *Protein Eng.* (2000), 13(2) 79-88.
178. Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S. & Thornton, J. M. PDBsum: Structural summaries of PDB entries. *Protein Sci.* (2018), 27(1), 129-134.
179. Yang, J., Roy, A. & Zhang, Y. Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics* (2013), 29, 2588-2595.
180. Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J. & Yaguchi, M. Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. *Protein Sci.* (1994), 3(3), 467-475.
181. Pollet, A. *et al.* Mutagenesis and subsite mapping underpin the importance for substrate specificity of the aglycon subsites of glycoside hydrolase family 11 xylanases. *Biochim. Biophys. Acta - Proteins Proteomics* (2010), 1804, 977-985.
182. Pettersen, E. F. *et al.* UCSF Chimera, A visualization system for exploratory research and analysis. *J. Comput. Chem.* (2004), 25(13), 1605-12.
-

183. Fersht, A. R. & Serrano, L. Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* (1993), 3(1), 75-83.
184. Kumar, P. R., Eswaramoorthy, S., Vithayathil, P. J. & Viswamitra, M. A. The tertiary structure at 1.59 Å resolution and the proposed amino acid sequence of a family-11 xylanase from the thermophilic fungus *Paecilomyces varioti* Bainier. *J. Mol. Biol.* (2000), 295, 581-593.
185. Tenkanen, M., Puls, J. & Poutanen, K. Two major xylanases of *Trichoderma reesei*. *Enzyme Microb. Technol.* (1992), 14(7), 566-574.
186. Scheiner, S., Kar, T. & Pattanayak, J. Comparison of various types of hydrogen bonds involving aromatic amino acids. *J. Am. Chem. Soc.* (2002), 124(44), 13257-13264.
187. Kannan, N. & Vishveshwara, S. Aromatic clusters: A determinant of thermal stability of thermophilic proteins. *Protein Eng.* (2000), 13(11), 753-761.
188. Tang, F. *et al.* Improving the thermostability of *Trichoderma reesei* xylanase 2 by introducing disulfide bonds. *Electron. J. Biotechnol.* (2017), 26, 52-59.
189. Harris, G. W. *et al.* Structural basis of the properties of an industrially relevant thermophilic xylanase. *Proteins Struct. Funct. Genet.* (1997), 29(1), 77-86.
190. Graubner, S., *et al.* New Xylanase with improved thermostability and increased enzyme activity on arabinoxylan. Patent. (2016), 24.
191. Arase, A. *et al.* Stabilization of xylanase by random mutagenesis. *FEBS Lett.* (1993), 316(2), 123-127.
192. Georis, J. *et al.* An additional aromatic interaction improves the thermostability and thermophilicity of a mesophilic family 11 xylanase: structural basis and molecular study. *Protein Sci* (2000), 9(3), 466-475.
193. Shibuya, H., Kaneko, S. & Hayashi, K. A single amino acid substitution enhances the catalytic activity of family 11 xylanase at alkaline pH. *Biosci. Biotechnol. Biochem.* (2005), 69, 1492-1497.
194. Zhang, S. *et al.* Five mutations in N-terminus confer thermostability on mesophilic xylanase. *Biochem. Biophys. Res. Commun.* (2010), 395, 200-206.
195. Wakarchuk, W. W. *et al.* Thermostabilization of the *Bacillus circulans* xylanase by

- the introduction of disulfide bonds. *Protein Eng. Des. Sel.* (1994), 7(11), 1379-1386.
196. Fenel, F., Leisola, M., Jänis, J. & Turunen, O. A de novo designed N-terminal disulphide bridge stabilizes the *Trichoderma reesei* endo-1,4- β -xylanase II. *J. Biotechnol.* (2004), 108, 137-143.
197. White, T. A. *et al.* CrystFEL: A software suite for snapshot serial crystallography. *J. Appl. Crystallogr.* (2012), 45(2), 335-341.
198. Daniel, R. M., Dines, M. & Petach, H. H. The denaturation and degradation of stable enzymes at high temperatures. *Biochemical Journal* (1996), 317(1), 1-11.
199. Lyubarev, A. E. & Kurganov, B. I. Analysis of DSC data relating to proteins undergoing irreversible thermal denaturation. *J. Therm. Anal. Calorim.* (2000), 62(1), 51-62.
200. Sanchez-Ruiz, J. M. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys. J.* (1992), 61(4), 921-935.
201. Ito, K., Ikemasu, T. & Ishikawa, T. Cloning and Sequencing of the xynA Gene Encoding Xylanase A of *Aspergillus kawachii*. *Biosci. Biotechnol. Biochem.* (1992), 56, 906-912.
202. Nakamura, S., Wakabayashi, K., Nakai, R., Aono, R. & Horikoshi, K. Purification and some properties of an alkaline xylanase from alkaliphilic *Bacillus* sp. strain 41M-1. *Appl. Environ. Microbiol.* (1993), 59(7), 2311-2316.
203. Törrönen, A., Harkki, A. & Rouvinen, J. Three-dimensional structure of endo-1,4- β -xylanase II from *Trichoderma reesei*: Two conformational states in the active site. *EMBO J.* (1994), 13, 2493-2501.
204. Sapag, A. *et al.* The endoxylanases from family 11: Computer analysis of protein sequences reveals important structural and phylogenetic relationships. *J. Biotechnol.* (2002), 95(2), 109-131.
205. Fushinobu, S., Ito, K., Konno, M., Wakagi, T. & Matsuzawa, H. Crystallographic and mutational analyses of an extremely acidophilic and acid-stable xylanase: Biased distribution of acidic residues and importance of Asp37 for catalysis at low pH. *Protein Eng.* (1998), 11, 1121-1128.
206. Tina, K. G., Bhadra, R. & Srinivasan, N. PIC: Protein Interactions Calculator.

- Nucleic Acids Res.* (2007), 35, 473-476.
207. Umemoto, H. *et al.* Contribution of salt bridges to alkaliphily of *Bacillus* alkaline xylanase. *Nucleic Acids Symp. Ser. (Oxf)*. (2007), 51, 461-462.
208. De Lemos Esteves, F. Improving the alkalophilic performances of the Xyl1 xylanase from *Streptomyces* sp. S38: Structural comparison and mutational analysis. *Protein Sci.* (2005), 14(2), 292-302.
209. Coolbear, T., Whittaker, J. M. & Daniel, R. M. The effect of metal ions on the activity and thermostability of the extracellular proteinase from a thermophilic *Bacillus*, strain EA.1. *Biochem. J.* (1992), 287(2), 367-374.
210. Pereira, J. de C. *et al.* Effect of metal ions, chemical agents and organic compounds on lignocellulolytic enzymes activities. *Enzyme Inhibitors and Activators* (2017), 29, 139-164.
211. Gao, J., Weng, H., Xi, Y., Zhu, D. & Han, S. Purification and characterization of a novel endo- β -1,4-glucanase from the thermoacidophilic *Aspergillus terreus*. *Biotechnol. Lett.* (2008), 30(2), 323-327.
212. Liao, H. *et al.* A new acidophilic endo- β -1,4-xylanase from *Penicillium oxalicum*: Cloning, purification, and insights into the influence of metal ions on xylanase activity. *J. Ind. Microbiol. Biotechnol.* (2014), 41(7), 1071-1083.
213. Liao, H. *et al.* Functional diversity and properties of multiple xylanases from *Penicillium oxalicum* GZ-2. *Sci. Rep.* (2015), 5, 1-14.
214. Zhang, F. *et al.* Cloning, expression, and characterization of an alkaline thermostable GH11 xylanase from *Thermobifida halotolerans* YIM 90462T. *J. Ind. Microbiol. Biotechnol.* (2012), 39, 1109-1116.
215. Debeche, T., Cummings, N., Connerton, I., Debeire, P. & O'Donohue, M. J. Genetic and biochemical characterization of a highly thermostable α -L- arabinofuranosidase from *Thermobacillus xylanilyticus*. *Appl. Environ. Microbiol.* (2000), 66, 1734-1736.
216. Nelson, D. L. & Cox, M. M. Princípios de bioquímica de Lehninger. *In Princípios de bioquímica de Lehninger* (2011), 975-975.
217. Zhang, F. *et al.* An acid and highly thermostable xylanase from *Phialophora* sp. G5.

- Appl. Microbiol. Biotechnol.* (2011), 89(6), 1851-1858.
218. Hu, J. & Saddler, J. N. Why does GH10 xylanase have better performance than GH11 xylanase for the deconstruction of pretreated biomass? *Biomass and Bioenergy* (2018), 110, 13-16.
219. Dobrev, G. *et al.* Purification and characterization of endoxylanase Xln-1 from *Aspergillus niger* B03. *World J. Microbiol. Biotechnol.* (2009), 25(12), 2095-2102.
220. Nair, S. G., Sindhu, R. & Shashidhar, S. Purification and biochemical characterization of two xylanases from *Aspergillus sydowii* SBS 45. *Appl. Biochem. Biotechnol.* (2008), 149(3), 229-243.
221. Vandermarliere, E. *et al.* Crystallographic analysis shows substrate binding at the -3 to +1 active-site subsites and at the surface of glycoside hydrolase family 11 endo-1,4- β -xylanases. *Biochem. J.* (2008), 410, 71-79.
222. Fowler, C. A. *et al.* Structure and function of a glycoside hydrolase family 8 endoxylanase from *Teredinibacter turnerae*. *Acta Crystallogr. Sect. D Struct. Biol.* (2018), 74, 946-955.
223. Moure, A., Gullón, P., Domínguez, H. & Parajó, J. C. Advances in the manufacture, purification and applications of xylo-oligosaccharides as food additives and nutraceuticals. *Process Biochem.* (2006), 41, 1913-1923.
224. Hughes, R. & Rowland, I. R. Stimulation of apoptosis by two prebiotic chicory fructans in the rat colon. *Carcinogenesis* (2001), 22(1), 43-47.
225. Gallaher, D. D. & Khil, J. The effect of synbiotics on colon carcinogenesis in rats. *Journal of Nutrition* (1999), 129(7), 1483-1487.
226. Gupta, P. K., Agrawal, P., Hedge, P. & Akhtar, M. S. Xylooligosaccharides and their anticancer potential: An update. *Anticancer Plants Nat. Prod. Biotechnol. Implements* (2018), 2, 255-271.
227. Aachary, A. A. & Prapulla, S. G. Xylooligosaccharides (XOS) as an emerging prebiotic: microbial synthesis, utilization, structural characterization, bioactive properties, and applications. *Compr. Rev. Food Sci. Food Saf.* (2011), 10(1), 2-16.
228. Fooks, L. J. & Gibson, G. R. In vitro investigations of the effect of probiotics and prebiotics on selected human intestinal pathogens. *FEMS Microbiol. Ecol.* (2002),

- 39(1), 67-75.
229. Yuan, X., Wang, J. & Yao, H. Antioxidant activity of feruloylated oligosaccharides from wheat bran. *Food Chem.* (2005), 90(4), 759-764.
230. Oku, T. & Nakamura, S. Digestion, absorption, fermentation, and metabolism of functional sugar substitutes and their available energy. *Pure Appl. Chem.* (2002), 74, 1253-1261.
231. Boukari, I., O'Donohue, M., Rémond, C. & Chabbert, B. Probing a family GH11 endo- β -1,4-xylanase inhibition mechanism by phenolic compounds: Role of functional phenolic groups. *J. Mol. Catal. B Enzym.* (2011), 72, 130-138.
232. Sharma, A., Milstein, O., Vered, Y., Gressel, J. & Flowers, H. M. Effects of aromatic compounds on hemicellulose-degrading enzymes in *Aspergillus japonicus*. *Biotechnol. Bioeng.* (1985), 27, 1095-1101.
233. Silva, C. de O. G. *et al.* GH11 xylanase from *Emericella nidulans* with low sensitivity to inhibition by ethanol and lignocellulose-derived phenolic compounds. *FEMS Microbiol. Lett.* (2015), 362(13), 94.
234. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* (2007), 2(9), 2212.
235. Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J. & Weiss, M. S. Optimization of protein buffer cocktails using Thermofluor. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* (2013), 69, 209-214.
236. Lo, M. C. *et al.* Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. *Anal. Biochem.* (2004), 332(1), 153-159.
237. Duan, D. *et al.* Internal structure and preferential protein binding of colloidal aggregates. *ACS Chem. Biol.* (2017), 12, 282-290.
238. Schöning-Stierand, K. *et al.* ProteinsPlus: interactive analysis of protein-ligand binding interfaces. *Nucleic Acids Res.* (2020), 48, 48-53.
239. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* (1995), 8(2), 127-134.

240. McGovern, S. L., Caselli, E., Grigorieff, N. & Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* (2002), 45(8), 1712-1722.
241. Alexander Ross, J. B., Laws, W. R., Rousslang, K. W. & Wyssbrod, H. R. Tyrosine fluorescence and phosphorescence from proteins and polypeptides. in *Topics in Fluorescence Spectroscopy* (2006).
242. Ghisaidoobe, A. B. T. & Chung, S. J. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on Förster resonance energy transfer techniques. *International Journal of Molecular Sciences* (2014).
243. Lakowicz, J. R. *Principles of fluorescence spectroscopy: Introduction to Fluorescence*, Springer, Boston, MA. (2006), 1-64.
244. Sopkova, J., Gally, J., Vincent, M., Lewit-Bentley, A. & Pancoska, P. The Dynamic Behavior of Annexin V as a Function of Calcium Ion Binding: A Circular Dichroism, UV Absorption, and Steady-State and Time-Resolved Fluorescence Study. *Biochemistry* (1994), 33(15), 4490-4499.
245. Gehlen, M. H. The centenary of the Stern-Volmer equation of fluorescence quenching: From the single line plot to the SV quenching map. *J. Photochem. Photobiol. C Photochem. Rev.* (2020), 42, 100338.
246. Chan, F. T. S. *et al.* Protein amyloids develop an intrinsic fluorescence signature during aggregation. *Analyst* (2013), 138, 2156-2162.
247. Vivian, J. T. & Callis, P. R. Mechanisms of tryptophan fluorescence shifts in proteins. *Biophys. J.* (2001), 80(5), 2093-2109.
248. Van De Weert, M. & Stella, L. Fluorescence quenching and ligand binding: A critical discussion of a popular methodology. *J. Mol. Struct.* (2011), 998, 144-150.
249. Engleder, M. *et al.* Recombinant expression, purification and biochemical characterization of kievitone hydratase from *Nectria haematococca*. *PLoS One* (2018), 13, 1-19.
250. Jhamb, K. & Sahoo, D. K. Production of soluble recombinant proteins in *Escherichia coli*: Effects of process conditions and chaperone co-expression on cell growth and production of xylanase. *Bioresour. Technol.* (2012), 123, 135-143.

251. Juturu, V. & Wu, J. C. Microbial xylanases: Engineering, production and industrial applications. *Biotechnology Advances* (2012), 30(6), 1219-1227.

6. List of Figures

Fig. 1.1 Steps of bioethanol production	3
Fig. 1.2 The phylogenetic tree	4
Fig. 1.3 Parameters for the efficient formulation of a biomass-degrading enzymatic cocktail.....	5
Fig. 1.4 An overview of enzymes used in wine making process and their applications.....	9
Fig. 1.5 Enzymatic hydrolysis of lignocellulose biomass.....	11
Fig. 1.6 Deutsches Elektronen Synchrotron (DESY)	16
Fig. 1.7 Protein Data Bank structure depositions.....	17
Fig. 1.8 Conventional vs serial crystallography.....	19
Fig. 1.9 Evolution of deposited PDB structures of GH11 xylanases (1994-2021).....	20
Fig. 1.10 Proposed catalytic mechanism of GH11 xylanases	21
Fig. 1.11 Inhibitors of GH11 enzymes.....	23
Fig. 2.1 Vector for <i>E.coli</i> expression of NhGH11	29
Fig. 2. 2 Schematic diagram of affinity chromatography	33
Fig. 2.3 Schematic diagram of soaking.....	43
Fig. 3.1 Summary of high entropy scores by SERp server	50
Fig. 3.2 Cleavage of MBP-NhGH11 with TEV protease	51
Fig. 3.3 Purification and physical characterization of NhGH11.....	52
Fig. 3.4 Crystallization of NhGH11	53
Fig. 3.5 Average B-factors and thermal ellipsoids.....	56
Fig. 3.6 Unconventional conformations.....	57
Fig. 3.7 Overall structure of NhGH11	59
Fig. 3.8 Surface exposed aromatic amino acids of NhGH11.....	60
Fig. 3.9 Active site cleft of NhGH11	61
Fig. 3.10 Insights into hydrogen bonding in the active site cleft.	62
Fig. 3.11 Representation of binding clefts on the surface of NhGH11.....	63
Fig. 3.12 Docking of xylohexose to the substrate-binding cleft of NhGH11.	65
Fig. 3.13 Sequence alignment and 3D comparison.....	67
Fig. 3.14 Surface representation and comparison of spatially conserved active site residues	68

Fig. 3.15 Thermostability in terms of thermal unfolding of secondary structure.	70
Fig. 3.16 Thermostability parameters	72
Fig. 3. 17 The temperature-dependent stability and activity.....	73
Fig. 3.18 The tape drive experiments.....	76
Fig. 3.19 Summary of physical and chemical variables for co-crystallization experiments of NhGH11.	78
Fig. 3.20 Co-Crystallization with inhibitors and substrates.....	79
Fig. 3.21 Comparison of crystal packing	80
Fig. 3.22 Effect of temperature on enzyme activity of NhGH11	81
Fig. 3.23 Effect of pH on enzyme activity of NhGH11	83
Fig. 3.24 Structural basis for optimum pH.....	83
Fig. 3.25 Effect of metal ions on the activity of NhGH11.....	85
Fig. 3. 26 Activators and inhibitors of GH enzymes	86
Fig. 3.27 Effect of chemical reagents on the activity of NhGH11.....	87
Fig. 3.28 Chemical structures of Xylan substrates.....	88
Fig. 3. 29 Substrate specificity and kinetic parameters.....	89
Fig. 3.30 Identification of catalysis products	91
Fig. 3.31 The proposed mechanism of xylopentose hydrolysis	92
Fig. 3.32 Residual activity of NhGH11 in presence of lignin-derived inhibitors	93
Fig.3.33 Thermofluor assay	94
Fig. 3.34 Effect of phenolic compounds on the stability of NhGH11	95
Fig. 3.35 Destabilization effect in presence of phenolic compounds	96
Fig. 3.36 Predicted interactions of NhGH11 with phenolic compounds	97
Fig. 3.37 Effect of tryptophan positions on emission spectra.....	100
Fig. 3.38. Tryptophan fluorescence intensity of NhGH11	102
Fig. 3.39 Binding constants of phenolic compounds	103

7. List of Tables

Table 1.1 Commercially available xylanases.....	7
Table 1.2 Carbohydrate-Active enzyme (CAZy) for biofuel production	11
Table 2.1 Consumables	26
Table 2.2 Instruments.....	27
Table 2.3 Bacterial Strains	30
Table 2.4 Medium and Buffers	31
Table 2.5 Buffers used for purification.....	33
Table 2. 6 Buffers used for purification of TEV.....	35
Table 2.7 Buffers and solutions used for SDS-PAGE	37
Table 2.8 Summary of soaking experiments.....	43
Table 2.9 Buffers and substrate solutions for biochemical assays.....	47
Table 3. 1 Data-collection and refinement statistics	54
Table 3. 2 Prediction of consensus binding site residues of GH11 xylanase.....	63
Table 3.3 The root mean square deviation of individual residues (Å) in the active site cleft	69
Table 3.4 Thermostability factors of NhGH11 in comparison to thermophilic xylanases ..	71
Table 3.5 Room temperature data collection	77
Table 3.6 Comparison of kinetic values of NhGH11 with various substrates	88
Table 3.7 Description of predicted binding sites of NhGH11	98

8. Risks and safety statements

Chemicals used (GHS classification)

Compounds	CAS-No.	GHS	Hazard Statements	Precautionary Statements
Acrylamide 37%	79-06-1	GHS06, GHS08	H301, H312, H315, H317, H319, H332, H340, H350, H316f, H372	P201, P280, P301+310, P305+351+338, P308+313
Acetic acid	64-19-7	GHS02, GHS05	H226, H314	P280, P305+351+338, P310
Agar	9002-18-0	-	-	-
Agarose	9012-36-6	GHS02	H226	P210, P240, P280, P303+261+353, P403+235, P501
Amylose	9005-82-7	-	-	-
Ammoniumsulfate	7783-20-2			
Ampicillin	69-52-3	GHS08	H334, H317	P280, P261, P302+352, P342+311
APS	7727-54-0	GHS03, GHS07, GHS08	H272, H302, H315, H317, H319, H334, H335	P280, P305+351+338, P302+352, P304+341, P342+311
Bromophenolblue	115-39-9	-	-	-
Calcium chloride	10043-52-4	GHS07	H319	P305+351+338
Caffeic acid	331-39-5	GHS07, GHS08	H315, H319, H335, H351, H361	P201, P202, P261, P264, P271, P280, P281, P302+352, P304+340, P305+351+338,

8. Risks and safety statements

				P308+313, P312, P321, P33 2+313, P337+313, P362, P403+233, P405, P501
Cinnamic acid	140-10-3	GHS07	H315, H319, H335	P261, P264, P271, P280, P302+352, P304+340, P305+351+338, P312, P321, P332+313, P337+313, P362, P403+233, P405, P501
Chloramphenicol	56-75-7	GHS08	H350	P201-P308+ P313
Cobalt chlorid	7646-79-9	GHS07, GHS08, GHS09	H350i, H360F, H302, H317, H334, H341, H410	P201, P273, P280- P302+352, P304+340, P342+311
Coomassie Brilliant Blue	6124-59-2	-	-	-
Coumaric acid	501-98-4	GHS05, GHS07, GHS08	H301, H302, H311, H314, H315, H317, H319, H335	P260, P261, P264, P270, P271, P272, P280, P301+310, P301+312, P301+330+331, P302+352, P303+361+353, P304+340, P305+351+338, P310, P312, P321, P322, P330,

8. Risks and safety statements

Copper chlorid	7447-39-4	GHS05, GHS06, GHS07, GHS09	H301, H302, H312, H315, H318, H319, H335, H400, H410, H411	P261, P264, P270, P271, P273, P280, P301+310, P301+312, P302+352, P304+340, P305+351+338, P310, P312, P321, P322, P330, P332+313, P337+313, P362, P363, P391, P403+233, P405,
DTT	3483-12-3	GHS07	H302, H315, H319, H335	P302,352, P305+351+338
3,5-Dinitrosalicylic acid				
Ethanol	64-17-5	GHS02	H225, H319	P210, P240, P305+351+338, P403+233
EDTA	60-00-4	GHS07	H319	P305+351+338
Ferulic acid	1135-24-6	-	-	-
Ferrous sulfate	7720-78-7	GHS07	H302, H315, H319	P305+351+338
Glucose	50-99-7	-	-	-
Glycerine	56-81-5	-	-	-
Glycine	56-40-6	-	-	-
Imidazole	288-32-4	GHS05, GHS07, GHS08	H302, H314, H360D	P201, P280, P301+P330+P331, P305+P351+P338, P308+P310
IPTG	367-93-1	-	-	-

8. Risks and safety statements

Isopropanol	67-63-0	GHS02, GHS07	H225, H319, H336	P210, P233, P240- P305+P351+P338, P403+P235
Phenol				
Potassium chloride	7447-40-7	-	-	-
Potassiumdihydrogenp hosphate	7778-77-0	-	-	-
Coper sulfate	7758-98-7	GHS07, GHS09	H302, H315, H319, H410	P273, P302+352, P305+351+338
Magnesium chlorid	7791-18-6	-	-	-
Manganese chloride	7773-01-5	-	-	-
Magnesium sulfate	7487-88-9	-	-	-
Maltose				
Manganese chloride	7773-01-5	GHS07	H302	P264, P301+312, P330, P501
Methnol		GHS02, GHS06, GHS08	H225, H301, H311, H331, H370	P210, P233, P240, P241, P242, P243, P260, P264, P270, P271, P280, P301+330+331, P310, P302+352, P312
β-Mercaptoethanol	60-24-2	GHS05, GHS06, GHS08, GHS09	H301+H331, H310, H315, H317, H318, H373, H410	P273, P280, P302+P352, P304+P340, P305+P351+P338, P308+P310
Lactose	63-42-3	-	-	-
Sodium azide	26628-22-8		H300, H310, H400, H410	P260, P280, P301+310, P501
Sodium chloride	7647-14-5	-	-	-
Sodium citrate	6132-04-3	-	-	-
Sodium	10049-21-5	-	-	-

8. Risks and safety statements

dihydrogenphosphate				
Sodium fluorid	7681-49-4	GHS06	H301, H315, H319	P302+352, P305+351+338, P308+310
Sodium sulfate	7757-82-6	-	-	-
Ni-NTA-Agarose	60-24-2	-	-	-
Nickel sulfat	10101-97-0	GHS07, GHS08, GHS09	H302, H332, H315, H317, H334, H341, H350i, H360D, H372, H410	P201, P273, P280, P302+352, P304+340, P308+313
PEG 400	25322-68-3	-	-	-
PEG 6000	25322-68-3	-	-	-
Hydrochloric acid	7647-01-0	GHS05, GHS07	H314, H335	P261, P280, P310, P305+351+338
SDS	151-21-3	GHS02, GHS06	H228, H302, H311, H315, H319, H335	P210, P261, P280, P301+312+330, P305+351+338+3 10, P370+P378
TEMED	110-18-9	GHS02, GHS05, GHS07	H225, H302, H314, H332	P261, P280, P305+351+338
Tris	1185-53-1	GHS07	H315, H319, H335	P261, P305+351+338
Triton-X100	9002-93-1	GHS05 GHS09 GHS07	H302,H315,H 318	
Tween-80	9005-65-6	-	-	-
Xylan beechwood	9014-63-5	-	-	-
Xylan polysaccharide	XYL100	-	-	-
Wheat arabinoxylm	9040-27-1	-	-	-

8. Risks and safety statements

Azo- xyloglucan	37294-28-3			
Xylose	58-86-6	-	-	-
Xylobiose	6860-47-5	-	-	-
Xylotriose	47592-59-6	-	-	-
Xylopentose	49694-20-4	-	-	-
Yeast extract	8013-01-2	-	-	-
Zinc chlorid	7646-85-7	GHS05, GHS07, GHS09	H302, H314, H410	P273, P280, P301+330+331, P305+351+338, P308+310

GHS pictograms



Hazard statements form GHS (Globally Harmonized System) and labels of chemicals

Symbol	GHS Hazard Statements
H225	Highly flammable liquid and vapor
H226	Flammable liquid and vapor
H228	Flammable solid
H272	May intensify fire; oxidizer
H301	Toxic if swallowed
H302	Harmful if swallowed
H311	Toxic in contact with skin
H312	Harmful in contact with skin
H314	Causes severe skin burns and eye damage
H315	Causes skin irritation
H317	May cause an allergic skin reaction
H318	Causes serious eye damage
H319	Causes serious eye irritation

8. Risks and safety statements

H330	Fatal if inhaled
H331	Toxic if inhaled
H332	Harmful if inhaled
H334	May cause allergy or asthma symptoms or breathing difficulties if inhaled
H335	May cause respiratory irritation
H336	May cause drowsiness or dizziness
H340	May cause genetic defects
H341	Suspected of causing genetic defects
H350	May cause cancer
H350i	May cause cancer by inhalation
H360	May damage fertility or the unborn child
H360D	May damage the unborn child
H360Fd	May damage fertility. Suspected of damaging the unborn child
H360FD	May damage fertility. May damage the unborn child
H361	Suspected of damaging fertility or the unborn child
H361d	Suspected of damaging the unborn child
H361f	Suspected of damaging fertility
H370	Cause damage to organs
H372	Causes damage to organs through prolonged or repeated exposure
H373	May cause damage to organs through prolonged or repeated exposure
H410	Very toxic to aquatic life with long lasting effects
H411	Toxic to aquatic life with long lasting effects
H412	Harmful to aquatic life with long lasting effects.

9. Acknowledgment

The increasing demand for fuels in parallel to the upsurge in population growth has generated considerable energy crises in Pakistan. Currently, the oil reserves of Pakistan are depleting and there is a need to find out sustainable and cleaner energy sources. These problems motivated me to find out the potential of agricultural lignocellulosic biomass for the production of biofuels. I am expecting to apply the knowledge of this research work for the development of biofuel industries in my country.

I am thankful to the Higher Education Commission, Pakistan, for funding of scholarship under the faculty development program of Bahauddin Zakariya University (BZU), Multan, Pakistan. Also, I am thankful to Ahmad Imran from the planning and development department, BZU, for organizing the scholarship-related issues very professionally.

I would like to express gratitude to my supervisor and mentor Prof. Dr. Dr. Christian Betzel for providing me this wonderful opportunity to work in his group. I want to pay appreciation to him for providing all facilities and infrastructure with highly equipped laboratories at University of Hamburg (UHH) and DESY campus.

I would say special thanks to Dr. Hévila Brognaro for her support in designing this project and biochemical studies. She guided me during the first steps of my Ph.D. studies. Also, her scientific proficiencies in biofuel research provided me consultation throughout my research work and enabled me to achieve my goals.

I would like to thank Prof. Dr. Wilhelm Schäfer for identifying *Nectria haematococca* and Dr. Steffen Ostendorp from the Department of Botany, UHH, for allowing me to use the facility for inhibition studies. I would like to thank Dr. Maria Riedner from organic chemistry, UHH, for providing a mass spectrometry facility. Also, I am thankful to Dominik Oberthür from DESY, for providing me the opportunity to learn serial crystallography. Also, I appreciate the beamline scientists at P11 and P13, DESY, for their guidance during data collection.

Thanks to all my work colleagues from AG Betzel, especially Dr. Markus Perbandt, Najeeb Ullah, Dr. Vasundara Srinivasan, Dr. Nadine Werner, Dr. Sven Falke, Dr. Christina Schmidt, Sebastian Fuchs, Dr. Célestin Mudogo, Bruno Alves Franca and Mengying Wang, for being a part of a good team and providing me the opportunities for scientific discussion. I am thankful to my friend Sumera Yamin and landlady Helga Schüler for helping me not to be homesick in Hamburg.

I am eternally grateful to my mother for her prayers and also to my siblings and friends from Pakistan for their best wishes.

Last but not least, this work would not be accomplished without my husband's love and support as he carried all responsibilities of my seven months old daughter and sent me to Hamburg for doctoral studies. He exemplified and promoted women's empowerment in the comparatively conservative society of Pakistan.

Thanks to everyone who supported me to attain my goals.

10. Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine anderen als die angegebenen Hilfsmittel benutzt zu haben. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde. Ferner versichere ich, dass ich noch keine Promotionsversuche an anderen Universitäten unternommen habe.

Hamburg, 16.03.2021

Hina Andaleeb