FUNCTIONAL INSIGHTS AND MOLECULAR MODEL ANALYSES OF *Thermotoga maritima* XYLANASES REVEAL THERMOSTABILITY AND COMPLEX EVOLUTIONARY LINEAGE

Aqsa Parvaiz[1], Muhammad Rizwan Javed[2], Ghulam Mustafa[1], Muhammad Aamer Mehmood[2], Muhammad Tahir Ul Qamar[3], Ayesha Pervaiz[4] and Faiz Ahmad Joyia[1,*]

[1]Center of Agricultural Biochemistry and Biotechnology (CABB), University of Agriculture, Faisalabad-38040, Pakistan; [2]Department of Bioinformatics and Biotechnology, Government College University, Faisalabad-38000, Pakistan; [3]College of Informatics, Huazhong Agricultural University, Wuhan-430070, P.R. China; [4]Institute of Pharmacy, Physiology and Pharmacology, University of Agriculture, Faisalabad-38040, Pakistan.
Corresponding author's e-mail: faizahmad1980@gmail.com

Thermotoga maritima derived xylanases are important industrial enzymes owing to extensive thermotolerance and stability under broad pH range This study presents *in-silico* characterization of *Thermotoga maritima* derived xylanase A and xylanase B through homology modeling and molecular docking approaches. Analyses of primary and secondary structures revealed net negative charges on both xylanase A and xylanase B. Modular structure of xylanase A exhibited five significant domains including two CBM49, one GlycoHydro10 and two DUF1083. Whereas, structure of xylanase B contained one significant domain i.e. Glyco hydro 10. The amino acid residues Leu-54, Arg-85, Leu-66, Phe-68, Asn-81 and Ile-52 of CBM49 have shown interaction with xylan during docking analyses. Within second CBM49 domain, Tyr-96, Trp-47 and Arg-19 have shown interaction with xylan. No residue of GlycoHydro10 was found to interact with xylan. Whereas, Asn-55, Arg-83, Arg-60 and ILe-118 of DUF1083 interacted with xylan. The Phe-170, Trp-156, Glu-15, Ser-157 and Lys-17 of DUF-1083 domain of xylanase A also showed interaction with xylan. The Glu-15 and Phe-170 of xylanase B showed H-bonding while Lys-565 showed electrostatic interaction with xylan. Comparative phylogenetic analysis of xylanases from *Thermotoga* genus, showed the complex evolutionary lineage of multimeric xylanase A protein. The findings of this study will lead us towards functional insights of xylanases from *Thermotoga maritima* and will help the researchers in xylanase enzyme engineering through rational design for developing highly efficient industrial enzymes.

Keywords: Xylanases, *Thermotoga maritima*, *in-silico* analyses, Homology modeling, molecular docking, substrate interaction

INTRODUCTION

Xylanases are ubiquitous enzymes produced by prokaryotes as well as eukaryotes (Meeta and Anil, 2013). Xylanses are involved in the hydrolysis of xylan which is an integral constituent of hemicellulosic cell wall (Kellete *et al.,* 1990; Black *et al.,* 1994). Xylan has great potential as renewable carbon and energy source through conversion to biofuels. Xylan is a hemicellulosic sub-component of phytobiomass that accounts for roughly 1/3rd of all the renewable organic carbon globally. Therefore, development of inexpensive technologies where hemicellulose can be used as primary substrate is essential (Alvira *et al.,* 2010). The xylanases of *Thermotoga maritima* are thermostable enzymes metabolizing a repertoire of carbohydrates including cellulose, hemicellulose, pectin, glucose, sucrose, xylan and many others (Nelson *et al.,* 1999). Recently, xylanases have appealed substantial research interest owing to impending applications in paper & pulp industry, food & feed industry and fermentation industry (Kumasaka *et al.,* 2005; Verma and

Satyanarayana, 2012). Xylanases hydrolyze xylans to yield mainly xylobiose as end product. Therefore, these enzymes are being employed in cost effective xylobiose production through xylan hydrolysis at commercial scale (Huang *et al.,* 2019). In paper and pulp industry, xylanases are replacing chlorine for pulp bleaching as hydrolysis of xylan removes lignin from plant-based pulp. Similarly, xylanases have potential applications in enzyme-driven processes like clarification of juices, improvement in beer consistency, enhancing digestibility of animal feed stock and above all conversion of lignocellulosic biomass and agro-wastes to fermentation end-products (Gessesse and Gashe, 1997).

Xylanases have significance for industrial applications by virtue of their stability at high temperature and alkaline conditions (Kumasaka *et al.*, 2005). The major advantage of using thermostable enzymes in industry is the enhanced rate of reaction at high temperature. For instance, the rate of reaction is known to be doubled by every 10°C rise in temperature, which in turn decreases the amount of enzyme

needed, subsequently decreasing the cost of production (Zhou *et al.,* 2011).

For the proper exploitation of these enzymes on industrial scale, detailed molecular characterization is essential. Although the protein structural and functional features can only be determined precisely by sophisticated experimental analyses (e.g. x-ray crystallography, NMR spectroscopy etc.), nonetheless; these analyses are very expensive, labor intensive and require larger amounts of purified sample as well as sound technical expertise (Ul Qamar and Khan, 2017). Computational tools are frequently available and are used for protein characterization and sequence analyses. These are cost effective and provide information nearly exact to be obtained via aforementioned experimental methods. Structural, structure-function relationship, interaction with substrate and other physico-chemical properties can be better explained using computational tools. However, comprehensive knowledge of structural biology and cutting-edge softwares are prime prerequisites for protein modeling (Sehar *et al.,* 2013a).

In the present study, Xylanase A and Xylanase B of *T. maritima* were characterized in an *in-silico* study. Xylanases were characterized on the basis of deduced 3D structures, physico-chemical properties and phylogenetic relationships. Later, the xylanase-xylan interactions were determined using molecular docking. The findings of this study will lead towards functional insights of xylanases from *Thermotoga maritima* and will help the researchers in xylanase engineering through rational design for future exploitation in industrial and environmental applications.

METHODOLOGY

Protein sequence analyses: Amino acids sequences of xylanase A and xylanase B were retrieved from Uniprot (Bairoch *et al.,* 2005) with accession numbers Q60037 and Q9WXS5, respectively. Primary sequence analyses of xylanases were performed using structure analyses tools available at ExPASy (Gasteiger *et al.,* 2005). The GRAVY index was determined using ProtParam tool available at ExPASy. Secondary structure analyses were performed using PSIPRED server (McGuffin *et al.,* 2000) and conserved domains were analysed using Pfam (Finn *et al.,* 2016).

Construction of 3D structures and their analyses: The 3D structure of xylanase B was available in PDB (ID:1VBR) while 3D structures of xylanase A domains were predicted using homology modeling approach as described previously (Sehar *et al.,* 2013b). Domain based structure modelling was performed because xylanase A exhibited exceptionally unique structure which could not be modeled using conventional protocols, and no suitable template was available for complete protein homology modeling. Domain wise multiple sequence alignment of xylanase A was performed by ClustalX2 (Ul Qamar and Khan, 2017)

Modeller 9.11 was employed for homology modelling and selection of template was the first step for homology modelling. For template selection, PSI-BLAST was run against PDB and 3D structure of templates were acquired from protein data bank (PDB). The template having at least 25% sequence identity and 60% query coverage was selected as per selection criteria (Fischer *et al.,* 2001). After template selection, the alignment between template and query sequence was performed using Align2D script of Modeller 9.11. Followed by the construction of 3D structure of different domains of xylanase A using ModellingScript.

Five (5) models for each domain were constructed and one of each with best quality (having higher values on verify 3D, quality factor, Ramachandran plot) was selected. The predicted 3D structure was subjected to MolProbity and NIH server (SAVES) for analyses. Through MolProbity Ramachandran values, poor rotamers, C_β deviations, bad bonds and bad angels of residues were computed (Chen *et al.,* 2010). Various evaluating tools including PROCHECK (Laskowski *et al.,* 1996), ERRAT (Colovos and Yeates, 1993) and Verify_3D (Eisenberg *et al.,* 1997) were accessed by NIH server (SAVES). The model for each domain with best values for all parameters (having higher values on verify 3D, quality factor, Ramachandran plot) was selected. Later on, 3D structure of template and query protein was superimposed by UCSF Chimera for verification.

Molecular docking studies: To study the enzyme-substrate interactions, both enzymes were subjected to molecular docking analyses using a pentamer of xylan as ligand by AutoDockVina (Gessesse and Gashe, 2011). AutoDockVina was used for flexible ligand-protein docking. All steps for docking were performed iteratively for each domain. The 3D structure of ligand (xylan pentamer) was retrieved from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The dock sites within protein structure were found by submitting the 3D structure of query protein at Pocket-Finder server (http://www.modelling.leeds.ac.uk/pocketfinder/). Different grid boxes were drawn for different domains. For studying molecular interaction of xylan with CBM49 domain of xylanase A, grid box having coordinates; centre-x = 95.497, centre-y = 24.935, centre-z = 31.206, size_x = 40, size_y = 40 and size_z = 40 was set. To find molecular interaction of xylan with second CBM49 domain of xylanase A, grid box having coordinates; centre-x = 36.303, centre-y = 25.102, centre-z = 8.919, size_x = 40, size_y = 40 and size_z = 40 was set. For molecular interaction of xylan with GlycoHydro10 domain of xylanase A, grid box having coordinates; centre-x = -50.736, centre-y = -20.285, centre-z = -94.894, size_x = 60, size_y = 54 and size_z = 40 was set. To determine molecular interaction of xylan with DUF1083 domain of xylanase A, grid box having coordinates; centre-x = 32.748, centre-y = 62.193, centre-z = 21.419, size_x = 40, size_y = 40 and size_z = 40 was set. For studying molecular interaction of xylan with second DUF1083 domain of xylanase A, grid box

having coordinates; centre-x = 27.353, centre-y = 43.946, centre-z = 16.382, size_x = 40, size_y = 40 and size_z = 40 was set. To determine interaction of xylan with xylanase B, a grid box having coordinates; centre-x = 22.644, centre-y = 14.977, centre-z =18.992, size_x = 40, size_y = 40 and size_z = 40 was set. Polar hydrogen molecules were added to all protein molecules and then docking analysis was carried out.

***Phylogenetic analysis*:** The evolutionary relationship of xylanase A and B from different bacteria was determined by reconstructing a phylogenetic tree using MEGA 7.0 (Hedges *et al.,* 2015). Neighbour-joining method was used to infer the evolutionary history, while the Poisson correction method was used to figure out the evolutionary distances and their unit was the number of amino acid substitutions per site.

RESULTS AND DISCUSSION

***Stability and solubility analyses*:** Here we have described structure prediction, important domains of xylanase A, molecular docking and phylogenetic analyses of xylanase A and xylanase B from *T. maritima.* We constructed 3D structure of xylanase A domains for the first time, refined and submitted to PDB (Table 1). According to our findings Xylanase A consisted of 273 residues. Among them, 154 residues showed negative while 119 residues appeared to have positive charges on them. On the other hand, xylanase B was relatively smaller molecule with a total of 96 residues with 52 negatively charged and 44 positively charged residues. Hence, it is concluded that both enzymes carry a net negative charge. Previously a thermostable family XynB (10 Endo-Xylanases) from *T. maritima* was characterized that cuts p-Nitrophenyl-β-D-xyloside (Zhengqiang *et al.,* 2001). High affinity binding was observed in thermostable xylanase A with both insoluble xylan and insoluble cellulose (Boraston *et al.,* 2001).

In the current study, physico-chemical properties like isoelectric point, instability index, GRAVY, amino acid and atomic compositions of the said enzyme were also determined. Summary of all such characteristics is given in Table 1. The instability indices of xylanase A & B were found to be 29.79 and 38.41, respectively. The stability of a query protein is estimated by its instability index. A protein with instability index less than 40 is predicted to be stable and vice

Table 1. Physico-Chemical Parameters of Xylanase A and Xylanase B (Determined by ProtParam).

Parameters	Xylanase A			Xylanase B		
Mol. Weight	119643.2 Dalton			40674.5 Dalton		
No. of amino acids	1059			347		
Theoretical *pI*	5.02			5.64		
Instability index (II)	29.79 (stable)			38.41 (stable)		
No. of Negatively Charged Residues (Asp + Glu)	154			52		
No. of Positively Charged Residues (Arg + Lys)	119			44		
Atomic Composition	Carbon	5407		Carbon	1867	
	Hydrogen	8340		Hydrogen	2836	
	Nitrogen	1386		Nitrogen	474	
	Oxygen	1638		Oxygen	525	
	Sulfur	21		Sulfur	11	
Amino Acid Composition	Ala (A)	67	6.3%	Ala (A)	20	5.8%
	Arg (R)	34	3.2%	Arg (R)	15	4.3%
	Asn (N)	54	5.1%	Asn (N)	20	5.8%
	Asp (D)	73	6.9%	Asp (D)	18	5.2%
	Cys (C)	1	0.1%	Cys (C)	4	1.2%
	Gln (Q)	36	3.4%	Gln (Q)	7	2.0%
	Glu (E)	81	7.6%	Glu (E)	34	9.8%
	Gly (G)	69	6.5%	Gly (G)	18	5.2%
	His (H)	12	1.1%	His (H)	8	2.3%
	Ile (I)	74	7.0%	Ile (I)	28	8.1%
	Leu (L)	69	6.5%	Leu (L)	27	7.8%
	Lys (K)	85	8.0%	Lys (K)	29	8.4%
	Met (M)	20	1.9%	Met (M)	7	2.0%
	Phe (F)	45	4.2%	Phe (F)	18	5.2%
	Pro (P)	46	4.3%	Pro (P)	14	4.0%
	Ser (S)	68	6.4%	Ser (S)	16	4.6%
	Thr (T)	68	6.4%	Thr (T)	11	3.2%
	Trp (W)	26	2.5%	Trp (W)	10	2.9%

versa (Guruprasad *et al.,* 1990). The present study revealed that xylanase A is more stable with instability index (29.79) than that of xylanase B with instability index (38.41). Second physiochemical parameter determined in present study was aliphatic index. The aliphatic index is termed as the relative volume of a protein molecule occupied by aliphatic side chains including alanine, valine, isoleucine and leucine. Aliphatic indices for xylanase A & B were found to be 82.81 and 87.64, respectively. High aliphatic index is an indication of thermal stability of globular proteins (Ikai, 1980). Hence, it is concluded that both the xylanase molecules under study are thermostable.

The Grand average of hydropathy (GRAVY) value for a protein molecule is computed by dividing sum of hydropathy values of all residues by total number of residues in that sequence. It helps in determining hydrophilic or hydrophobic nature of a protein molecule. Hydrophilic nature of protein molecules was indicated by negative GRAVY value (Kyte and Doolittle, 1982). GRAVY value for xylanase A was -0.390 and for xylanase B was -0.367 which showed that both molecules were charge polarized and capable of establishing hydrogen bonds due to which their interaction with water and other polar solvents was expected to be favourable thermodynamically. Hence, this study revealed that both the enzymes under study were thermostable as well as hydrophilic in nature.

***Secondary structure and domain analyses*:** The secondary structures of xylanase A and B were analyzed which exhibited 4.2% α helices, 32.86% β sheets and 62.94% loops in xylanase A while xylanase B contained 35% α helices, 11% β sheets and 54% loops. Domain analysis revealed that both enzymes possess multidomain structures which is a direct clue of a complex evolutionary track of said enzymes. Xylanase A contained five significant and two insignificant domains. Significant domains included two CBM49, one GlycoHydro10 and two DUF1083 while insignificant domains were DUF2374 and Laminin G3. It has been

described that domains are sequences and structures of proteins that remained conserved during the course of evolution. Such a multi-domain structure of xylanase A reflects that the enzyme may have role in diverse metabolic pathways other than xylan degradation (Teng *et al.,* 2014). On the other hand, xylanase B showed one significant and one insignificant domain including GlycoHydro10 and AAA18, respectively (Table 2a, 2b). The multiple sequence alignment of *T. maritima* xylanases showed their association with other members of *Thermotoga* genus, particularly xylanase A which has shown a close resemblance with *Thermotoga* sp. EMP, while xylanase B showed close resemblance with *T. neapolitana* (Yang and Han, 2018). Some most recent studies also revealed that xylanases with multidomain structures are more diverse in functions and more resilient to extreme conditions especially temperatures (Teo *et al.,* 2019). Similarly, Nakamichi *et al*. (2019) also characterized a multidomain Xlyanase from a fungus *Talaromyces cellulolyticus* and found it to be bifunctional.

***Homology modeling and molecular docking*:** The 3D structure of xylanase B was retrieved from PDB (1VBR) while structures of various domains of xylanase were predicted through homology modelling using Modeller 9.11. The criterion of template selection was based on sequence identity greater than 25% and lower E-value using PSI-BLAST. For 1st CBM49 domain, 2W5F was selected as a template with 28% identity; For 2nd CBM49 domain, 1H6X was selected as a template with 33% identity; for GlycoHydro10 domain, 1N82 was selected as template with 46% identity; and for both DUF83 domains, 1I82 was used as a template with 26% and 100% identity, respectively. The best quality model is selected out of 5 for each domain of xylanase A (Fig. 1). Several methods including Ramachandran plot, Verify3D and ERRAT available at NIH server (SAVES) were used to check the consistency of predicted 3D structures (Table 3). Poor rotamers, C_β deviations, bad bonds and bad angels were determined

Table 2a. Domain analysis of Xylanase A (Determined by Pfam).

Domains	Description	Amino acid Position	PMDB Ids	Signifiance
CBM49 (1st)	Carbohydrate binding domain	50-184	PM0082205	Significant Pfam-A match
CBM49 (2nd)	Carbohydrate binding domain	203-340	PM0082206	Significant Pfam-A match
GlycoHydro10	Glycosyl hydrolase family 10	369-693	PM0082207	Significant Pfam-A match
DUF1083 (1st)	Domain of unknown Function	715-869	PM0082208	Significant Pfam-A match
DUF1083 (2nd)	Domain of unknown Function	882-1059	PM0082209	Significant Pfam-A match
DUF2374	Protein of unknown function	14-41		Insignificant Pfam-A match
Laminin G3	Concanavalin A-like lectin/glucanases superfamily	238-367		Insignificant Pfam-A match

Table 2b. Domain analysis of Xylanase B (Determined by Pfam).

Domains	Description	Amino acid Position	Significance
GlycoHydro10	Glycosyl hydrolase family 10	24-339	Significant Pfam-A match
AAA 18	AAA domain	24-113	Insignificant Pfam-A match

through MolProbity (Table 4). All final models passed the quality filtering tests. The 3D structures were furthermore validated by superimposition of predicted structures with their respective templates. UCSF Chimera's structure comparison command for superimposition was used for this purpose (Fig. 2), and it was assured that predicted models are accurate. Xylan binds to xylanases to change its structure due to hydrolysis and increases its activity.
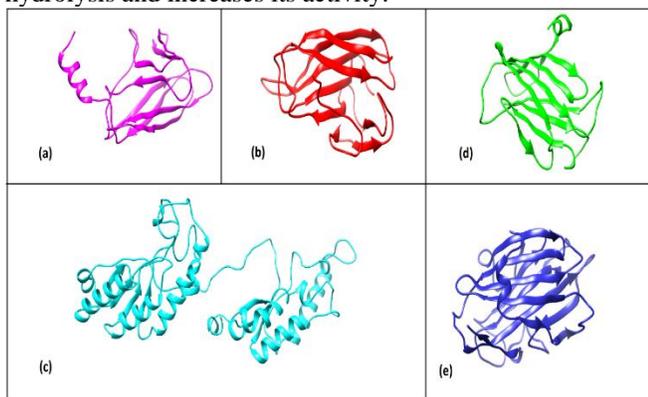
Figure 1. **Models of Xylanase A domains, visualized by UCSF Chimera. (a) 1st CBM49 domain (b) 2nd CBM49 domain (C) GlycoHydro10 domain (d) 1st DUF83 domain (e) 2nd DUF83 domain.**

The interactions do not cause inactivation of xylanases, so the hydrolysis rate is increased (Kaya *et al.,* 2000). When ligand binds to receptor (enzyme) molecule, it alters the function of enzyme. The binding affinity can be characterized by the interaction of most ligands with their binding sites on receptors. High-affinity ligand binding is characterized by greater intermolecular force between ligand and receptors and vice versa (Eldridge *et al.,* 1997).

Table 3. NIH Server (SAVES) analyses of template and models of the Xylanase A predicted by MODELLER. NIH Server (SAVES) evaluates the best selected models by using tools like ERRAT, Verify_3D and PROCHECK.

Models	Verify3D	ERRAT Quality Factor	PROCHECK Ramachandran
2WYS (Template)	97.07%	97.811	89.0%
Xylanase A (1st CBM49)	71.32%	30.952	89.8%
1H6X (Template)	90.63%	84.768	86.8%
Xylanase A (2nd CBM49)	99.28%	31.538	91.7%
1N82 (Template)	89.98%	98.148	91.1%
Xylanase A (GlycoHydro10)	85.09%	59.105	91.1%
1I82 (Template)	97.89%	86.188	88.0%
Xylanase A (1st DUF1083)	74.02%	44.915	88.2%
1I82 (Template)	97.89%	86.188	88.0%
Xylanase A (2nd DUF1083)	91.01%	88.757	94.2%

**Table 4. MolProbity analyses of template and models of Xylanase A predicted by MODELLER.**

Models	Poor Rotamers	Ramachandran Outliers	Ramachandran Favored	Cβ deviations >0.25Å	Residues with bad bonds	Residues with bad angles
2WYS (Template)	2.32%	0.79%	95.47%	0.31%	0.02%	0.06%
Xylanase A (1st CBM49)	2.56%	1.50%	91.73%	0.00%	0.00%	0.00%
1H6X (Template)	0.78%	0.00%	98.09%	0.00%	0.00%	0.00%
Xylanase A (2nd CBM49)	2.54%	1.47%	97.79%	0.00%	0.00%	0.00%
1N82 (Template)	1.59%	0.00%	98.32%	0.15%	0.00%	0.06%
Xylanase A (GlycoHydro10)	2.45%	1.88%	95.30%	0.65%	0.00%	0.81%
1I82 (Template)	0.63%	0.00%	96.26%	0.00%	0.00%	0.32%
XylanaseA (1st DUF 1083)	0.89%	2.42%	92.74%	0.84%	0.00%	0.16%
1I82 (Template)	0.63%	0.00%	96.26%	0.00%	0.00%	0.32%
Xylanase A (2nd DUF 1083)	2.60%	0.00%	98.29%	0.00%	0.00%	0.00%
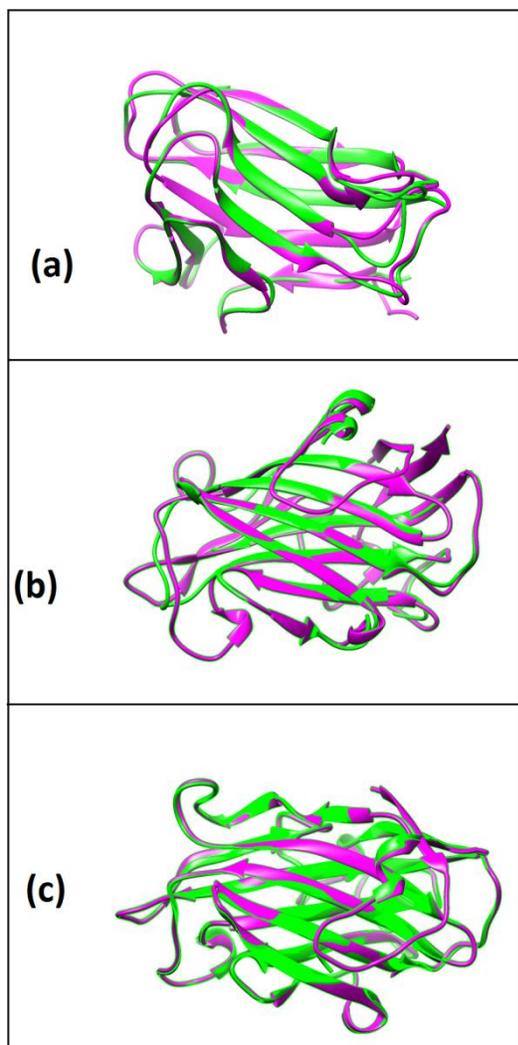
Figure 2. **Superimposition of models with their entire template, using UCSF Chimera. (a) Superimposition of CBM49 domain (Green) with 1H6X template (Magenta), (b) Superimposition of DUF83 domain (Green) with 1I82 template (Magenta), (c) Superimposition of DUF83 domain (Green) with 1I82 template (Magenta).**

After docking, interaction analyses were performed. The amino acid residues Leu-54, Arg-85, Leu-66, Phe-68, Asn-81 and Ile-52 of CBM49's domain 1 have shown interaction with xylan. Among them, The Asn-81, Ile-52, Leu-54 and Arg-85 showed H-bonding with the substrate. For second CBM49 domain, Tyr-96, Trp-47 and Arg-19 showed interaction with xylan. No residue of GlycoHydro10 was observed to bind with xylan. For DUF1083, Phe-120, Asn-55, Arg-83, Arg-60 and ILe-118 of DUF1083 interacted with xylan. Phe-120, Asn-55 and Arg-83 showed hydrogen bonding, while Arg-60

and Ile-118 showed electrostatic interactions. In xylanase A fifth domain, Phe-170, Trp-156, Glu-15, Ser-157 and Lys-17 showed interaction with xylan. In case of xylanase B, Glu-15 and Phe-170 showed H-bonding, while Lys-565 showed electrostatic interactions with xylan (Fig. 3). Binding specificity of CBM9-2, a domain of xylanase A, has already been reported. Binding of CBM9-2 to barley xyloglucan and β-glucan and weakly to wheat arabinoxylan and soluble birchwood xylan endorsed the findings of present study. The receptors for hydroxyethylcellulose, dextran arabinan, arabinogalactan, CM-cellulose, pecticgalactan, starch, laminarin, or locust bean gum are not present. The direct interaction of CBM9-2 with reducing end of polysaccharides is the most novel property (Yang and Han, 2018; Teo *et al.,* 2019; Tassa *et al.,* 2009). Xylanase B could bind to both insoluble xylan and Avicel even though it had no activity on Avicel. Several solvent-exposed aromatic rings mediate the binding by hydrogen bonding and stacking interactions with sugar in polysaccharides (Huang *et al.,* 2019; Nakamichi *et al.,* 2019).
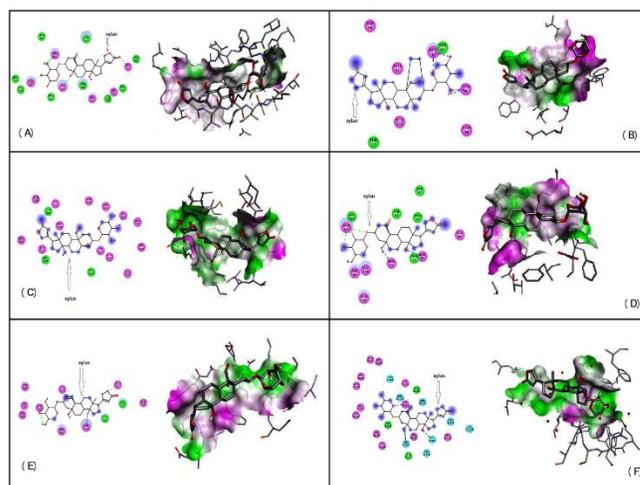
Figure 3. **Residues of all domains of Xylanase A and Xylanase B interacting with xylan are shown. (A) 1st CBM49 domain interaction with xylan, (B) 2nd CBM49 domain interaction with xylan, (C) GlycoHydro10 domain interaction with xylan, (D) 1st DUF83 domain interaction with xylan, (E) 2nd DUF83 domain interaction with xylan, (F) GlycoHydro10 domain of xylanase B interaction with xylan.**

Phylogenetic analyses: The xylanases have been sequenced from a variety of bacteria and sequences have been submitted to protein sequence databases. Evolutionary relationship
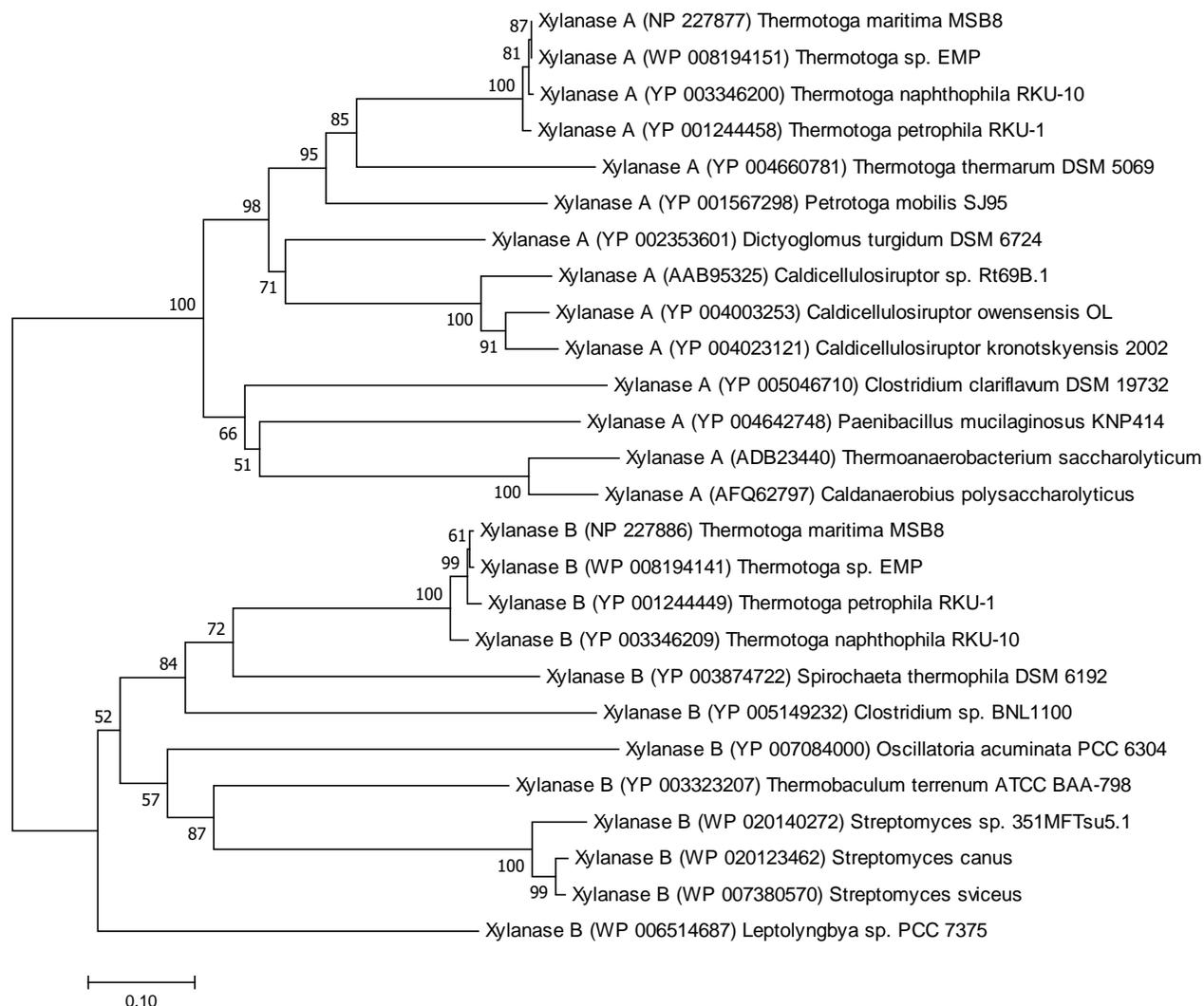
Figure 4. Phylogenetic tree based on microbial xylanase amino acid sequences.

among submitted sequences can be revealed by sequences alignment and subsequent phylogenetic analysis of xylanase A & B from a variety of microorganisms. The phylogenetic analyses showed that xylanase A and B of *T. maritima* have diverse evolutionary relationship, not only within the genus Thermotoga but also with other microbial genera. Both xylanases were grouped in different clans showing their evolutionary divergence. We used neighbour-joining method to deduce evolutionary history of enzymes under study (Ul Qamar and Khan, 2017). Figure 4 shows optimal tree with a total of branch length = 5.27560916. The percentage of replicate trees where the related taxa clustered together in the bootstrap test (1000 replicates) was marked next to the branches. The units of branch lengths in phylogenetic tree were same as those of the evolutionary distances used while inferring the phylogenetic tree. The scale bar corresponds to

0.1-estimated amino acid substitution per sequence position. The analysis involved 26 amino acid sequences with elimination of gaps and missing data. Final dataset contains total of 302 positions. MEGA7 software package was used to conduct evolutionary analyses (Sehar *et al.,* 2013b). Xylanase amino acid sequences along with accession numbers (NCBI) from different microorganisms are represented in taxa. This further demonstrates that variety of physiochemical and thermodynamic characteristics of xylanases isolated from different groups of organisms also represents variation at protein sequence level.

*Epilogue***:** This study leads us to better understanding of enzyme-substrate interaction in prokaryotes especially in extremophiles. It also lead us towards functional insights of xylanases from *Thermotoga maritima* and will help the researchers interested in xylanase enzyme engineering

through rational design to develop highly efficient enzymes for future exploitation in industrial and environmental applications.

REFERENCES

Alvira, P., E. Tomás-Pejó, M. Ballesteros and M.J. Negro. 2010. Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review. Biores. Technol. 101:4851-4861.

Bairoch, A.M., R. Apweiler, C.H. Wu, W.C. Barker, B. Boeckmann, S.F. Rojas, E. Gasteiger, H. Huang, R. Lopez, M. Magrane and M.J. Martin. 2005. The universal protein resource (UniProt). Nucl. Acids Res. 33:D154-D159.

Black, G.W., G.P. Hazlewood, G.P. Xue, C.G. Orpin and H.J. Gilbert. 1994. Xylanase B from *Neocallimastix patriciarum* contains a non-catalytic 455-residue linker sequence comprised of 57 repeats of an octapeptide. Biochem. J. 299:381-387.

Boraston, A.B., A.L. Creagh, M.M. Alam, J.M. Kormos, P. Tomme, C.A. Haynes, R.A. Warren and D.G. Kilburn. 2001. Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. Biochemistry 40:6240-6247.

Chen, V.B., W.B. Arendall, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, L.W. Murray, J.S. Richardson and D.C. Richardson. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66:12-21.

Colovos, C. and T.O. Yeates. 1993. Verification of protein structures: patterns of nonbonded atomic interactions. Prot. Sci. 2:1511-1519.

Eisenberg, D., R. Lüthy and J.U. Bowie. 1997. VERIFY3D: Assessment of protein models with three-dimensional profiles. Meth. Enzymol. 277:396-404.

Eldridge, M.D., C.W. Murray, T.R. Auton, G.V. Paolini and R.P. Mee. 1997. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. J. Computer-aided Mol. Des. 11:425-445.

Finn, R.D., P, Coggill, R.Y. Eberhardt, S.R. Eddy, J. Mistry, A.L. Mitchell, S.C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas and G.A. Salazar. 2016. The Pfam protein families database: towards a more sustainable future. Nucl. Acids Res. 44:D279-D285.

Fischer, D., A. Elofsson, L. Rychlewski, F. Pazos, A. Valencia, B. Rost, A.R. Ortiz and R.L. Dunbrack. 2001. CAFASP2: the second critical assessment of fully automated structure prediction methods. Proteins 45:171-183.

Gasteiger, E., C. Hoogland, A. Gattiker, S.E. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch. 2005. Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook Humana Press.

Gessesse, A. and B.A. Gashe. 1997. Production of alkaline xylanase by an *alkaliphilic Bacillus* sp. isolated from an alkalinesoda lake. J. Appl. Microbiol. 83:402-406.

Guruprasad, K., B.V. Reddy and M.W. Pandit. 1990. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting *in vivo* stability of a protein from its primary sequence. Prot. Eng. Des. Sel. 4:155-161.

Hedges, S.B., J. Marin, M. Suleski, M. Paymer and S. Kumar. 2015. Tree of life reveals clock-like speciation and diversification. Mol. Biol. Evol. 32:835-845.

Huang, Y., X. Zheng, B. Pilgaard, J. Holck, J. Muschiol, S. Li and L. Lange. 2019. Identif ication and characterization of GH11 xylanase and GH43 xylosidase from the chytridiomycetous fungus, *Rhizophlyctis rosea*. Appl. Microbiol. Biotechnol. 103:777–791.

Ikai, A. 1980. Thermostability and aliphatic index of globular proteins. J. Biochem. 88:1895-1898.

Kaya, F., J.A. Heitmann and T.W. Joyce. 2000. Influence of lignin and its degradation products on enzymatic hydrolysis of xylan. J. Biotechnol. 80:241-247.

Kellett, L.E., D.M. Poole, L.M. Ferreira, A.J. Durrant, G.P. Hazlewood and H.J. Gilbert. 1990. Xylanase B and an *arabinofuranosidase* from *Pseudomonas fluorescens* subsp. cellulosa contain identical cellulose-binding domains and are encoded by adjacent genes. Biochem. J. 272:369-376.

Kumasaka, T., T. Kaneko, C. Morokuma, R. Yatsunami, T. Sato, S. Nakamura and N. Tanaka. 2005. Structural basis of the substrate subsite and the highly thermal stability of xylanase 10B from *Thermotoga maritima* MSB8. PROTEINS: Struct. Funct. Bioinform. 61:999-1009.

Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105-132.

Laskowski R.A., J.A. Rullmann, M.W. MacArthur, R. Kaptein and J.M. Thornton. 1996. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J. Biomol. NMR 8:477-486.

McGuffin, L.J., K. Bryson and D.T. Jones. 2000. The PSIPRED protein structure prediction server. Bioinformatics 16:404-405.

Meeta, S. and K. Anil. 2013. Xylanases: an overview. Brit. Biotechnol. J. 3:1-28.

Nakamichi, Y., T. Fouquet, S. Ito, M. Watanabe, A. Matsushika and H. Inoue. 2019. Structural and functional characterization of a bifunctional GH30-7 xylanase B from the filamentous fungus Talaromyces cellulolyticus. J. Biol. Chem. 294:4065-4078.

Nelson, K.E., R.A. Clayton, S.R. Gill, M.L. Gwinn, R.J. Dodson, D.H. Haft, E.K. Hickey, J.D. Peterson, W.C.

Nelson, K.A. Ketchum and L. McDonald. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323-329.

Sehar, U., M.A. Mehmood, K. Hussain, S. Nawaz, S. Nadeem, M.H. Siddique, H. Nadeem, M. Gull, N. Ahmad, I. Sohail and S.S. Gill. 2013a. Domain wise docking analyses of the modular chitin binding protein CBP50 from *Bacillus thuringiensis* serovar konkukian S4. Bioinformation 9:901-907.

Sehar, U., M.A. Mehmood, S. Nawaz, K. Hussain, I. Sohail, M.R. Tabassum, S.S. Gill and A. Saqib. 2013b Three dimensional (3D) structure prediction and substrate-protein interaction study of the chitin binding protein CBP24 from *B. thuringiensis*. Bioinformation 9:725-729.

Tassa, C., J.L. Duffner, T.A. Lewis, R. Weissleder, S.L. SchreiberL, A.N. Koehler and S.Y. Shaw. 2009. Binding affinity and kinetic analysis of targeted small molecule-modified nanoparticles. Bioconj. Chem. 21:14-19.

Teng, Z., M. Guo, Q. Dai, C. Wang, J. Li and X. Liu. 2014. Computational prediction of protein function based on weighted mapping of domains and GO terms. BioMed Res. Int. 2014:1-9.

Teo, S.C., K.J. Liew, M.S. Shamsir, C.S. Chong, N.C. Bruce, K.G. Chan and K.M. Goh. 2019. Characterizing a Halo-Tolerant GH10 Xylanase from *Roseithermus sacchariphilus* Strain RA and Its CBM-Truncated Variant. Int. J. Mol. Sci. 2019; 20:2284.

Ul Qamar M.T. and M.S. Khan. 2017. A novel structural and functional insight into chloroplast-encoded central subunit of dark-operated protochlorophyllide oxidoreductase (DPOR) of plants. Pak. J. Agri. Sci. 54:395-406.

Verma, D. and T. Satyanarayana. 2012. Molecular approaches for ameliorating microbial xylanases. Biores. Technol. 117:360-367.

Yang, J. and Z. Han. 2018. Understanding the Positional Binding and Substrate Interaction of a Highly Thermostable GH10 Xylanase from *Thermotoga maritima* by Molecular Docking. Biomolecules 8:64. doi:10.3390/biom8030064

Zhengqiang, J., A. Kayashi, M.M. Ahsan, L. Lite, M. Kitaoka and K. Hayashi. 2001. Characterization of a thermostable family 10 endo-xylanase (XynB) from *Thermotoga maritima* that cleaves p-nitrophenyl-β-D-xyloside. J. Biosc. Bioeng. 92:423-428.

Zhou, P., H. Zhu, Q. Yan, P. Katrolia and Z. Jiang. 2011. Purification and properties of a *psychrotrophic Trichoderma* sp. xylanase and its gene sequence. Appl. Biochem. Biotechnol. 164:944-956.