

# Determination of Substrate Specificities Against $\beta$ -Glucosidase A (BglA) from *Thermotoga maritima*: A Molecular Docking Approach

Muhammad Ibrahim Rajoka<sup>†</sup>, Sobia Idrees<sup>\*†</sup>, Usman Ali Ashfaq, Beenish Ehsan, and Asma Haq

Department of Bioinformatics and Biotechnology, Government College University (GCU), Faisalabad, Pakistan

Received: December 16, 2013  
Revised: June 12, 2014  
Accepted: August 29, 2014

First published online  
September 1, 2014

\*Corresponding author  
Phone: +923338988480;  
E-mail: sb.genny@gmail.com

<sup>†</sup>These authors contributed  
equally to this work.

pISSN 1017-7825, eISSN 1738-8872

Copyright© 2015 by  
The Korean Society for Microbiology  
and Biotechnology

Thermostable enzymes derived from *Thermotoga maritima* have attracted worldwide interest for their potential industrial applications. Structural analysis and docking studies were performed on *T. maritima*  $\beta$ -glucosidase enzyme with cellobiose and *p*NP-linked substrates. The 3D structure of the thermostable  $\beta$ -glucosidase was downloaded from the Protein Data Bank database. Substrates were downloaded from the PubCehm database and were minimized using MOE software. Docking of BglA and substrates was carried out using MOE software. After analyzing docked enzyme/substrate complexes, it was found that Glu residues were mainly involved in the reaction, and other important residues such as Asn, Ser, Tyr, Trp, and His were involved in hydrogen bonding with *p*NP-linked substrates. By determining the substrate recognition pattern, a more suitable  $\beta$ -glucosidase enzyme could be developed, enhancing its industrial potential.

**Keywords:** Active site residues,  $\beta$ -glucosidase, *Thermotoga maritima*, 3D modeling, molecular docking

## Introduction

The structure-function relationship and their extended half-life make enzymes from extremophiles preferred biocatalysts for industrial applications.  $\beta$ -Glucosidase (E.C. 3.2.1.21) is widely distributed in all kingdoms [5]. Generally speaking,  $\beta$ -glucosidases cleaves the  $\beta$ -1,4-glucosidic bonds in a variety of naturally occurring glucosides.  $\beta$ -Glucosidase, in synergism with other cellulases, converts cellobiose and heterosaccharides into glucose and removes feedback inhibition of exoglucanase (E.C. 3.2.1.31) and endoglucanase (E.C. 3.2.1.4) [13]. The available glucose can then be fermented into ethanol or methane.  $\beta$ -Glucosidase plays a role in cell wall functions like catabolism, animal metabolism of certain molecules, insect and microbe interaction with plants, plant aromas, activation of plant defense system, and release of plant hormones [2].

In humans, patients suffering from Gaucher's disease lack the genes for formation of membrane-bound lysosomal acidic  $\beta$ -glucosidase. These patients are administered intravenous injection of purified  $\beta$ -glucosidase from human placenta [2].  $\beta$ -Glucosidases have been used to hydrolyze

the flavanoids and isoflavanoid glucosides, found naturally in fruits, red wines, soyabeans, tea, and vegetables, releasing antitumor aglycones [13].

Various biocatalyst genes from thermophilic organisms have been cloned into mesophilic hosts and purified. Enormous efforts exploring the thermostability mechanisms of these enzymes at high temperatures have been performed [6, 8]. Genome structural analysis has revealed that 182 basic protein folds are present, with these folds performing diverse but related functions in most cases [14].  $\beta$ -Glucosidases are categorized into three classes depending upon substrate specificity [6].

Industrial application of  $\beta$ -glucosidase requires a large-scale production of the enzyme and detailed knowledge of its reaction mechanism. Understanding the mechanisms of  $\beta$ -glucosidase thermostability would provide insights and pave the way to design other hyperstable biocatalysts for industrial applications. Carefully selected amino acid sequence-based criteria distinguish them better than purely structure-based parameters. However, combined sequence and structure based indices may improve their performance somewhat further. From previous analysis, the strongest

contributors to thermostability are an enhancement in ion pairs on a protein's outer surface and a strong hydrophobic interior. Current research focuses on understanding the molecular basis of their substrate specificities and to catalog them into multifunctional biocatalysts. Identifying amino acid residues occurring at the enzyme active site and their vicinity is important for revealing the structure–function relationship, as well as designing mutant enzymes with improved thermal and hydrogen ion characteristics.  $\beta$ -Glycosidases formation with known substrate specificity has not yet been carried out in great detail. This is mainly because of the lack of data on the molecular mechanism of the  $\beta$ -glucosidase specificity; namely, data on the energy of the noncovalent interactions in the enzyme-transition state formation, and the stabilization process. The interaction determined by structural analysis and docking studies on  $\beta$ -glucosidase with natural substrate cellobiose (a hydrolytic product of cellulases) and all *p*-NP-linked substrates will lead to the catalysis of this important enzyme for its use in the biofuel production process.

## Materials and Methods

$\beta$ -Glucosidase activities can be measured based on (i) glucose released from a natural substrate, cellobiose [15], and (ii) fluorescence or colored products released from analogs of cellobiose; that is, *p*-nitrophenyl-beta-D-1,4-glucopyranoside (*p*-NPG), 4-methylumbelliferyl-beta-D-glucopyranoside

(salicin) using a glucose kit (HUMAN, Germany).

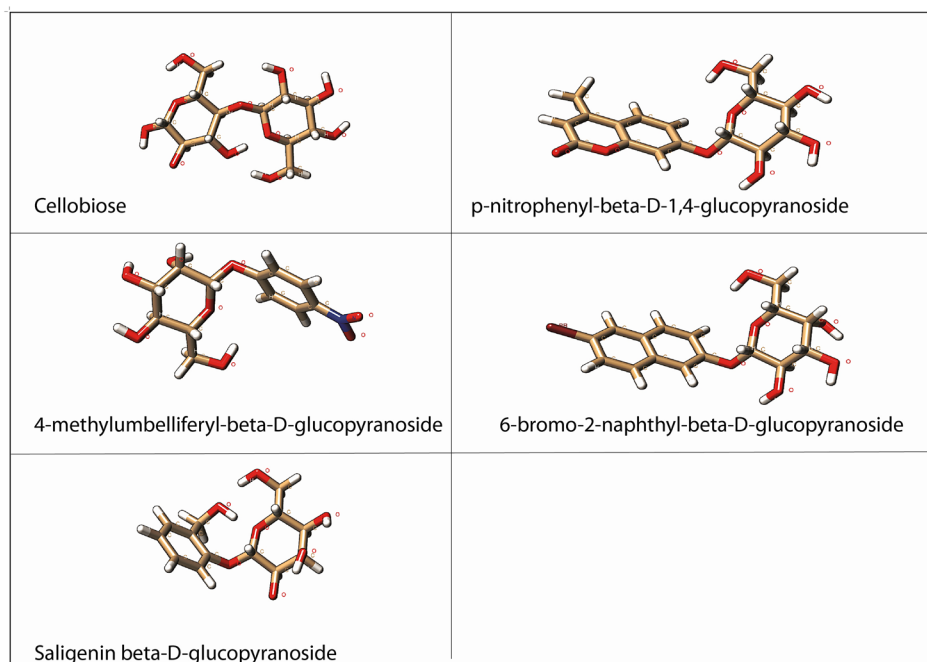
### Molecular Docking

The 3D structure of the BglA was downloaded from the Protein Data Bank (PDB) using PDB ID: 2WC4. 3D protonation of BglA was carried out using the Molecular Operating Environment (MOE) tool. The protein structure was minimized using MMFF94X Forcefield. 3D Structures of the substrates cellobiose (molecular formula:  $C_{12}H_{22}O_{11}$ ), *p*-nitrophenyl- $\beta$ -D-1,4-glucopyranoside (molecular formula:  $C_{12}H_{15}NO_8$ ), 4-methylumbelliferyl-beta- $\beta$ -D-glucopyranoside (molecular formula:  $C_{16}H_{18}O_8$ ), 6-bromo-2-naphthyl- $\beta$ -D-glucopyranoside (molecular formula:  $C_{16}H_{17}BrO_8$ ), and salicin (molecular formula:  $C_{13}H_{18}O_7$ ) were downloaded in standard delay format from the PubChem database, optimized by adding hydrogens, and removing water molecules. MMFF94S Forcefield was also used to 3D protonate and energy minimize the substrates profile. MOE [9] was applied for enzyme substrate docking. After protein docking, the resultant complexes were subjected to hydrogen bonding analysis. The LigX feature of MOE was used to find interactions among the ligand and receptor protein.

## Results

### Molecular Docking

The 3D structure of BglA was downloaded from the PDB database. Substrates (Fig. 1) were downloaded in standard



**Fig. 1.** 3D structures of cellobiose, *p*-nitrophenyl- $\beta$ -D-1,4-glucopyranoside, 4-methylumbelliferyl- $\beta$ -D-glucopyranoside, 6-bromo-2-naphthyl- $\beta$ -D-glucopyranoside, and saligenin  $\beta$ -D-glucopyranoside.

delay format and converted to 3D PDB files. After preparing receptor (enzyme) and ligands (substrates), BglA and cellobiose were subjected to docking using MOE software. The best docking complex was used to analyze the interactions of enzyme and cellobiose. Fig. 2 shows docked complexes of BglA with different substrates, and Fig. 3 shows binding of substrates with the hydrophobic surface of the enzyme. Three hydrogen-bond interactions were obtained in the BglA-cellobiose complex, as shown in Fig. 3 cellobiose. This reaction involved residues Asn223, Ser229, and His298 making hydrogen bonds with the ligand, suggesting these residues may play as an essential role in the pathway of substrate enzyme hydrolysis. The surrounding amino acids Trp122, Glu166, Gly224, Tyr294, Tyr295, Trp324, Phe414, and Glu405 (Fig. 2 cellobiose and Fig. 3) provide a hydrophobic environment enhancing thermostability.

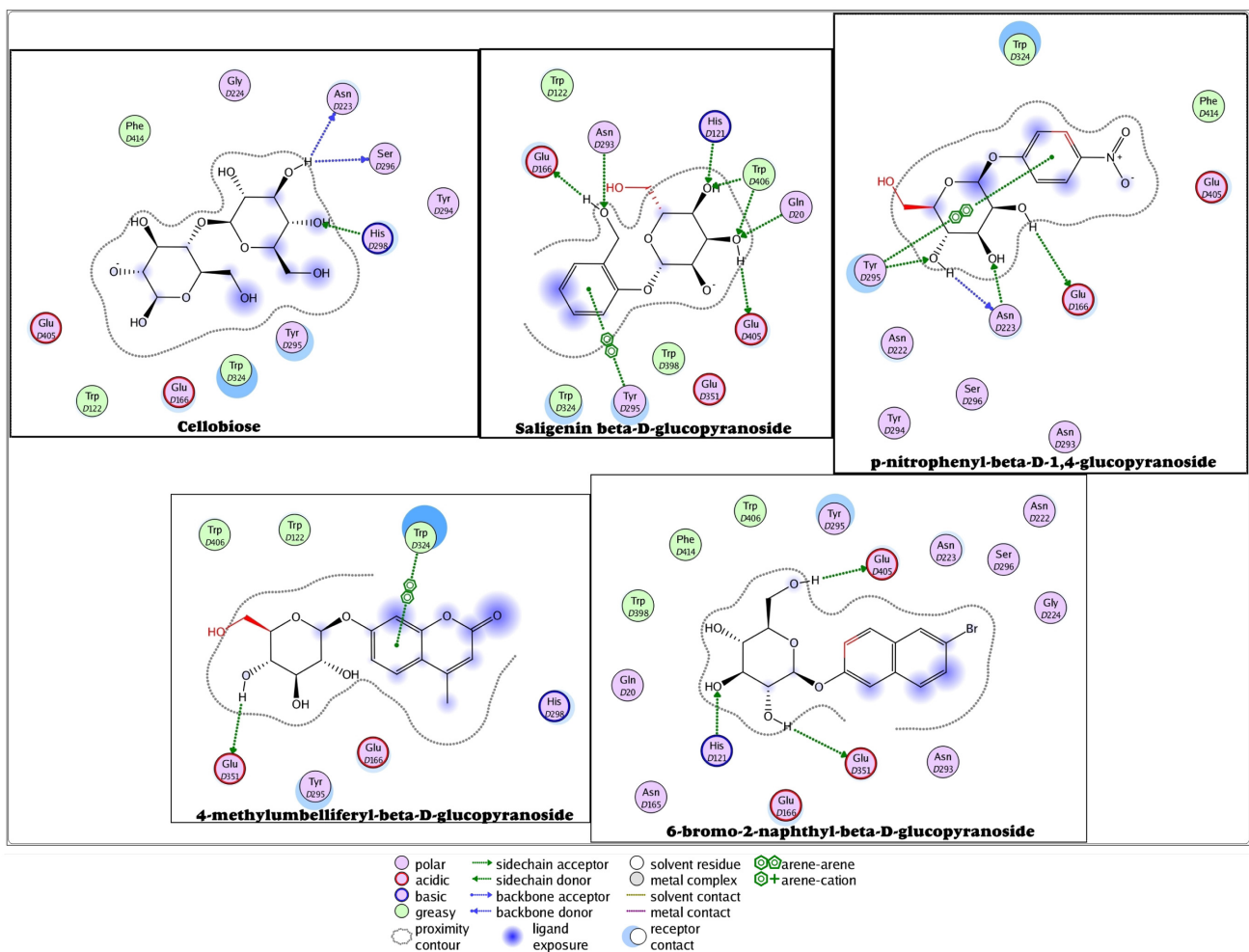
Saligenin  $\beta$ -D-glucoside (salicin) molecules were bonded with the following active site residues: Gln20, His121,

Glu166, Tyr295, Asn298, Trp406, and Glu405. Surrounding amino acids Trp122, Glu351, Trp324, and Trp398 (Fig. 2 salicin and Fig. 3) provided a hydrophobic environment for enhanced thermostability, as mentioned above.

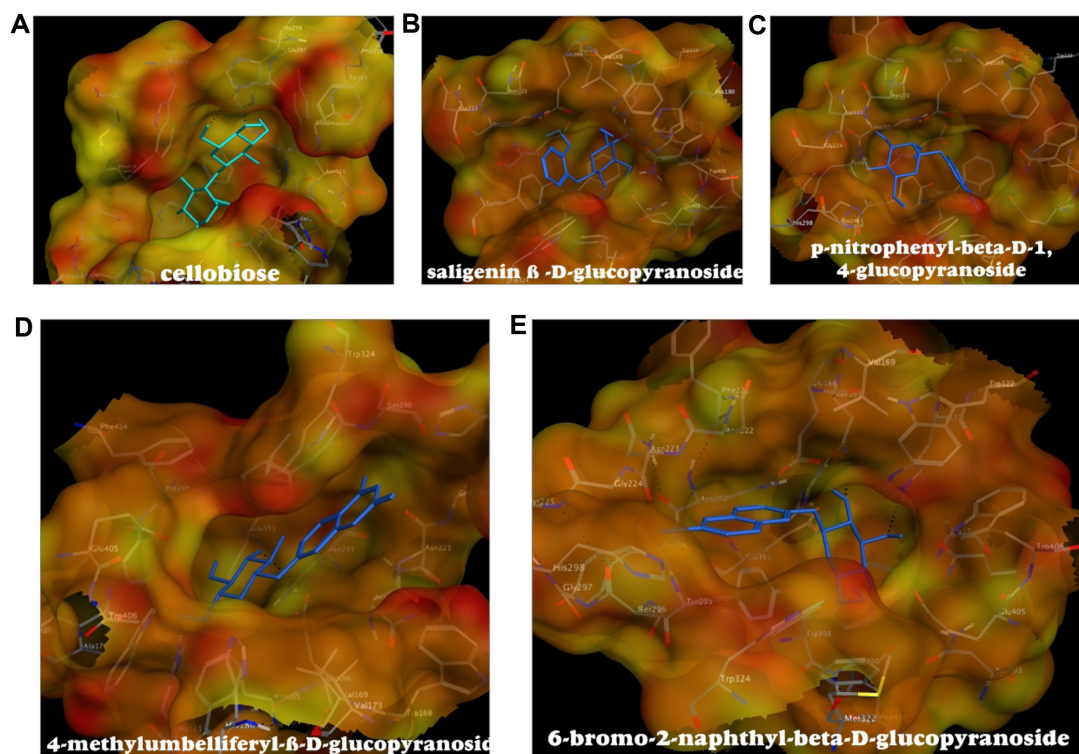
The *p*-nitrophenyl- $\beta$ -D-glucopyranoside molecules in the system were coordinated with the following active site residues: Glu166, Tyr295, and Asn223. This interacting system was surrounded by Asn222, Asn293, Tyr294, Ser296, Trp324, Glu405, and Phe414 (Fig. 2 *p*-nitrophenyl- $\beta$ -D-glucopyranoside and Fig. 3), respectively involved in imparting hydrophobicity to the enzyme substrate system.

The 4-methylumbelliferyl- $\beta$ -D-glucopyranoside (MUG) molecules in the system were coordinately linked with the following active site residues: Trp324 and Glu351. A hydrophobic environment for enhanced reactivity and thermophily was provided by Trp122, Glu166, Tyr295, His298, and Trp406 (Fig. 2 MUG and Fig. 3) respectively.

The 6-bromo-2-naphthyl- $\beta$ -D-glucopyranoside (BNG)



**Fig. 2.** Interaction analysis of  $\beta$ -glucosidase with substrates using the MOE LigX feature.



**Fig. 3.** Binding of *p*-nitrophenyl sugars with 3D surface of the BglA enzyme.

(A) Binding of cellobiose with BglA. (B) Binding of *p*-nitrophenyl-beta-D-1,4-glucopyranoside with BglA. (C) Binding of 4-methylumbelliferyl-beta-D-glucopyranoside with BglA. (D) Binding of 6-bromo-2-naphthyl-beta-D-glucopyranoside with BglA. (E) Binding of Saligenin beta-D-glucopyranoside with BglA.

molecules in the system were linked with the following active site residues: His121, Glu351, and Glu405. The hydrophobic environment for enhanced reactivity in this interaction was supported by Gln20, Asn222, Asn223, Gly224, Ser296, Asn293, Glu166, Asn165, Trp398, Trp406, and Phe414 (Fig. 2 BNG and Fig. 3), respectively. During temperature-induced agitational movement, Glu residues may get exchanged with Asn221 and may take part in active catalysis.

## Discussion

The catalytic mechanism of glycosidases involves two carboxylic acids: one acts as a general catalytic acid that donates protons to the *O*-glycosidic linkage, enhancing the leaving group's departure, while the other acts as a general catalytic base assisting in the nucleophilic attack on the carbon of the substrate. Another amino acid is involved in substrate bonding through hydrogen bonding [12]. Molecular interaction, especially hydrogen bonding, plays an important role in forming enzyme-substrate complexes. The glutamate sidechain can adapt its position to an axial

O4 and may have an ability to recognize glucoside-configured substrates due to its conformational freedom characteristics. The cellobiose was bonded with the polar residues Asn223 and Ser296, which were acting as backbone acceptors. The basic His298 also had receptor contact with the ligand and was acting as a sidechain donor.

Earlier, Haq and colleagues [4] using different docking models showed that BglA of *T. petrophila* contained two or three conserved Glu residues (Glu 166, 351, and 405) in close contact to the substrate to catalyze the reaction. In our studies, Glu166 and Glu405 were in hydrophobic contact with cellobiose and may have a role in recognition of inhibitors, as reported earlier by Mazlan and Khairudin [7]. Moreover, Trp122, Glu166, Gly224, Tyr294, Tyr295, Trp324, Phe414, and Glu405 (Fig. 2 cellobiose and Fig. 3) were present on the periphery of the enzyme-cellobiose complex. These provide a hydrophobic surrounding for enhanced thermostability and reactivity, as suggested earlier for thermophilic enzymes [1]. Thus, it can be inferred that the observed interactions are important in BglA and cellobiose binding and can contribute in catalyzing the enzyme action to produce glucose.



In salicin and BglA interaction, polar Gln20 and Asn293 were forming hydrogen bonds with the salicin and were acting as sidechain donors. The greasy residue Trp406 was working as a sidechain donor and had two hydrogen bonds. The acidic limbs of Glu166 and Glu405 were acting as sidechain acceptors and were having hydrogen bonding with the ligand. Similarly, Tyr295 was having strong receptor contact with the salicin. In its close proximity, the surrounding amino acids Trp122, Glu351, Trp324, and Trp398 (Fig. 2 salicin and Fig. 3) were also present to provide a hydrophobic environment for enhanced thermostability, as mentioned above. Hydrophobic interactions become stronger as the reaction temperature is enhanced, and increased thermal agitation at thermophilic temperatures prevents thermal unfolding. If hydrophobic entities are in close proximity to the enzyme-substrate complex, a negative entropy effect is frequently observed [1]. BglA is expected to undergo conformational change at thermophilic temperatures, which rearrange its structure in such a way as to enhance its hydrophobic interactions, as observed for xylanase from a *Scopulariopsis* sp. [1].

In the BglA-*p*-nitrophenyl- $\beta$ -D-glucopyranoside complex, the active site residue Tyr295 was having strong receptor contact as well as acting as a backbone donor. The Glu166 (acidic limb) was acting as a backbone acceptor. Polar residue Asn223 had two interactions, enabling it to work as a backbone donor and sidechain acceptor. Asn222, Asn293, Tyr294, Ser296, Trp324, Glu405, and Phe414 were present in close proximity to beta-glucoside. During temperature agitational movements, it may come closer to the substrate and may take part in active catalysis (in substrate bonding), as reported earlier by Haq and colleagues [4].

The MUG molecules in the system were coordinately linked through Trp324, which had arene-arene interaction with the ligand. Moreover, Glu351 (acidic limb) was acting as a backbone acceptor.

The hydrophobic environment for enhanced reactivity and thermophily was provided by Trp122, Glu166, Tyr295, His298, and Trp406 (Fig. 2 MUG and Fig. 3). As mentioned earlier [1], hydrophobic interactions are weakened at mesophilic temperatures but become stronger when the temperature is increased to thermophilic range to enhance its reactivity.

BNG activities were controlled by Glu351, Glu405 (acidic limb), and His121 (basic limb) as active site residues. The hydrophobic environment for enhanced reactivity in this interaction was supported by Gln20, Asn222, Asn223, Gly224, Ser296, Asn293, Glu166, Asn165, Trp398, Trp406, and Phe414 (Fig. 2 BNG and Fig. 3), respectively. Our

earlier studies on determining active site residues of  $\beta$ -glucosidase derived from *Cellulomonas biazotea* showed that  $pK_{a1}$  and  $pK_{a2}$  values of 5.5 and 7.9 complexed with values of heat of ionization of active site residues controlling the  $V_{max}$  of  $\beta$ -glucosidase had Glu or Asp as the acidic limb and His as the basic limb in catalyzing the hydrolysis of  $\beta$ -glucoside [11]. The Glu residue is dominantly found among family glycoside hydrolase 1.

Moreover, our data on enzyme-substrate docking studies also confirmed the reports of earlier workers [3] who experimentally proved the presence of Glu-based catalytic sites for BglA through enzyme inhibition studies. Tiwari and colleagues [12] through site-directed mutagenesis after altering Glu 445 with other amino acids confirmed that a conserved amino acid (Glu 445) in the substrate binding pocket was an essential residue affecting substrate binding affinity. In this way, they emphasized the active role of Glu along with other neighboring amino acids in catalyzing the hydrolysis of the substrate. This reflects that BglA is not only stable at elevated temperatures but can also catalyze the hydrolysis of *p*-NP-linked substrates as well. From the post docking analysis, it was found that Glu or acidic amino acid Asn were the main interacting residues with the enzyme. Haq and colleagues [4] also confirmed with their different docking models that highly thermostable BglA derived from *T. petrophila* contained two or three conserved Glu residues (Glu 166, 351, and 405) in close contact to the substrate to catalyze the reaction.

The substrate-enzyme complex structure in combination with mutagenesis analysis of the catalytic residues proposed the catalytic mechanism involved in enzyme catalysis. These studies highlight the involvement of a Glu-Glu catalytic mechanism. Current studies on the above lines have clearly indicated a new module that implicates a distinctive catalytic role of histidine, which may function as an intermediate for the electron transfer network between the typical Glu-Glu catalytic mechanism as reported by Haq *et al.* [4]. Further investigation suggested that the aromatic residue Trp as well as polar residues Asn, His, and Tyr in the active site not only participate in substrate binding but also provide a unique microenvironment suitable for active catalysis.

Application of Protein Ligand Interaction Fingerprints (PLIF) is a method of choice to identify the interacting residues between enzyme and ligands. The top ten ranked conformation complexes from each ligand (Fig. 3) were used as an input in the PLIF program to determine common active residues among the above-mentioned complexes involved in active catalysis. In agreement with other

workers [6], it is reported that Glu163, Glu349, Glu 403, and Tyr293 (along with others, see below) play an important role in catalytic function for BglA and ligand complexes.

In our study, after post docking analysis, it was found that BglA had strong interactions with all five substrates; thus, it can be inferred that  $\beta$ -glucosidase can catalyze the hydrolysis of the natural substrate cellobiose and the *p*-NP-linked substrates salicin, *p*-NPG, MUC, and BNG, corroborating the known information on this enzyme reported by other authors. The docking analysis resulted in identification of Glu, Asn Tyr, Trp, Ser, and His as important residues in protein-ligand interaction. The results from this study will eventually give new ideas to biologists to design a better enzyme for a more efficient enzymatic hydrolysis process of  $\beta$ -glucosides with lower production cost but higher yields.

## References

1. Afzal A, Ali S, Latif F, Rajoka M, Siddiqui K. 2005. Innovative kinetic and thermodynamic studies of a purified superactive xylanase from *Scopulariopsis* sp. *Appl. Biochem. Biotechnol.* **120**: 51-70.
2. Bhatia Y, Mishra S, Bisaria VS. 2002. Microbial beta-glucosidases: cloning, properties, and applications. *Crit. Rev. Biotechnol.* **22**: 375-407.
3. Gloster TM, Madsen R, Davies GJ. 2007. Structural basis for cyclohexitol inhibition of a beta-glucosidase. *Org. Biomol. Chem.* **5**: 444-446.
4. Haq I, Khan M, Muneer B, Hussain Z, Afzal S, Rashid N, et al. 2012. Cloning, characterization and molecular docking of a highly thermostable  $\beta$ -1,4-glucosidase from *Thermotoga petrophila*. *Biotechnol. Lett.* **34**: 1703-1709.
5. Ketudat Cairns JR, Esen A. 2010. Beta-glucosidases. *Cell Mol. Life Sci.* **67**: 3389-3405.
6. Koffi Y, Konan H, Kouadio J, Dabonne S, Due E, Kouame K. 2012. Purification and biochemical characterization of beta-glucosidase from cockroach. *J. Animal Plant Sci.* **13**: 1747-1757.
7. Mazlan N, Khairudin N. 2014. Docking study of  $\beta$ -glucosidase B (BglB) from *Paenibacillus polymyxa* with cellobiose and cellotetrose. *J. Med. Bioeng.* **3**: 78-83.
8. Myung S, Wanga Y, Zhang Y-H. 2010. Fructose-1,6-bisphosphatase from a hyper-thermophilic bacterium *Thermotoga maritima*: characterization, metabolite stability, and its implications. *Process Biochem.* **45**: 1882-1887.
9. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. 2005. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**: W363-W367.
10. Setlow B, Cabrera-Hernandez A, Cabrera-Martinez RM, Setlow P. 2004. Identification of aryl-phospho-beta-D-glucosidases in *Bacillus subtilis*. *Arch. Microbiol.* **181**: 60-67.
11. Siddiqui KS, Rashid MH, Rajoka MI. 1997. Kinetic analysis of the active site of an intracellular beta-glucosidase from *Cellulomonas biazotea*. *Folia Microbiol. (Praha)* **42**: 53-58.
12. Tiwari M, Lee K, Kalyani D, Singh R, Kim H, Lee J, Ramachandran P. 2012. Role of Glu445 in the substrate binding of  $\beta$ -glucosidase. *Process Biochem.* **47**: 2365-2372.
13. Xue Y, Song X, Yu J. 2009. Overexpression of  $\beta$ -glucosidase from *Thermotoga maritima* for the production of highly purified aglycone isoflavones from soy flour. *World J. Microbiol. Biotechnol.* **25**: 2165-2172.
14. Zhang Y, Thiele I, Weekes D, Li Z, Jaroszewski L, Ginalski K, et al. 2009. Three-dimensional structural view of the central metabolic network of *Thermotoga maritima*. *Science* **325**: 1544-1549.
15. Zhang YH, Lynd LR. 2006. A functionally based model for hydrolysis of cellulose by fungal cellulase. *Biotechnol. Bioeng.* **94**: 888-898.