

Molecular Modelling of a Thermostable Glycoside Hydrolase from *Caldivirga maquilingsensis* and Its Substrate Docking Mechanism for Galactooligosaccharides Synthesis

Rebaone Letsididi^{1,2}, Kekgabile S. Letsididi¹, Tao Zhang^{1,3}, Bo Jiang^{1,3}, Wanmeng Mu^{1,3,*}

¹State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi, Jiangsu, China

²National Food Technology Research Centre, Private Bag 008, Kanye, Botswana

³Synergetic Innovation Center of Food Safety and Nutrition, Jiangnan University, Wuxi, China

*Corresponding author: wmmu@jiangnan.edu.cn

Abstract Glycoside hydrolases (GHs) are very important enzymes that can catalyze the breakdown of the glycosidic bonds between carbohydrates and non-carbohydrates to synthesize GOS prebiotic sugars and hydrolyze lactose in the dairy industry. GOS can stimulate the growth of gut microbiota to have beneficial health effects in the host. The current work investigated molecular modelling and lactose substrate docking of a thermostable GH family 1 from *Caldivirga maquilingsensis* strain IC-167 and determined its mechanism for GOS synthesis. The 3D model structure obtained from the Swiss model analysis tool revealed that GH 1 from *Caldivirga maquilingsensis* was a tetramer with a catalytic pocket at the center of each monomer. The overall quality of the model was 93%. When lactose was docked in the catalytic site using AutoDock Vina software, the catalytic amino acid residues were identified to be Glu 387 and Glu 209 which acted as nucleophile and acid/base residues respectively. Other amino acid residues like His 153, Gln 19, Glu 432, Ala 266, Asn 267, Ser 268 and Trp 433 were also found to be surrounding the catalytic site and playing an essential role in ligand binding and recognition of the lactose substrate for GOS synthesis. These findings offer an understanding of how enzyme protein structure determines catalytic specificity, which serves as new knowledge basis to engineer GH 1 from *C. maquilingsensis* for the biosynthesis of GOS with a broad or narrow degree of polymerization range.

Keywords: molecular modeling, docking, glycoside hydrolase, thermostable, galacto-oligosaccharides, *Caldivirga maquilingsensis*

Cite This Article: Rebaone Letsididi, Kekgabile S. Letsididi, Tao Zhang, Bo Jiang, and Wanmeng Mu, "Molecular Modelling of a Thermostable Glycoside Hydrolase from *Caldivirga maquilingsensis* and its Substrate Docking Mechanism for Galactooligosaccharides Synthesis." *Biomedicine and Biotechnology*, vol. 6, no. 1 (2018): 1-7. doi: 10.12691/bb-6-1-1.

1. Introduction

Production of non-digestible galacto-oligosaccharides (GOS) has continued to gain importance because of their recognition as prebiotic compounds that stimulate growth of bifidobacteria and lactobacilli bacteria associated with beneficial health effects and also as an interesting functional food ingredient [1,2]. Furthermore, GOS are of special interest because they have structurally related oligosaccharides present in human breast milk [3,4]. Because of these benefits, GOS are of great interest for both human and animal nutrition.

Enzymatic synthesis of GOS from lactose is normally conducted using glycoside hydrolases (GH, EC 3.2.1.X) because these enzymes can transfer glycosyl moieties from a donor sugar to an acceptor [5,6]. This method is the one currently used in the dairy industry for GOS synthesis from lactose with the yield of the reaction

depending on the relative ratio of the transgalactosylation versus hydrolysis reaction [7,8]. In the reaction for lactose hydrolysis, a galactosyl-enzyme intermediate is formed and instead of transferring the galactose unit to water for hydrolysis reaction, the enzyme transfers it to any suitable nucleophile acceptor sugar present in the reaction system to synthesize various GOS. As a result, GOS are very complex mixtures consisting of numerous and different oligosaccharides varying in their degree of polymerization (DP), glycosidic linkages and anomeric configurations. In general terms, enzymatic hydrolysis of the glycosidic bond takes place via general acid catalysis that requires two critical residues: a proton donor and a nucleophile/base [9,10].

Glycoside hydrolases (GHs) with β -galactosidase activity occur in a variety of microorganisms from the super kingdoms bacteria, archaea and eukaryota. Some of these enzymes have been expressed in host organisms by several conventional techniques. Due to their superior stability properties, glycoside hydrolases from thermophilic

microorganisms have become an attractive alternative to the non-thermostable enzymes because they can allow high substrate concentrations to be used due to increased substrate solubility leading to increased GOS yields [11,12]. Reaction yield and specificity of β -galactosidases by the donor and acceptor of galactose depends on the enzyme origin due to differences in its structure and/or in the coupling mechanism of the sugars, leading to different distributions of products of synthesis [13,14]. Tertiary structure and amino acids in the active site affect β -galactosidase reaction specificity, being both important since the ability of the enzyme to accept nucleophiles other than water in its active center will determine the proper transfer of the galactose moiety [9]. Specificity will also be affected by the concentrations of the donor and acceptor of the transgalactosylated galactose [11,12].

The number of known protein sequences has currently increased following the publication of the human genome and other sequencing projects and deposited in the gene bank hence the increase in the number of protein sequences far exceeds the number of known protein structures [15]. As such the three dimensional (3D) structure modelling of enzymes is a very important tool used to reveal the identity of important amino acid residues involved in catalysis and to elucidate the mechanism in the enzymatic synthesis of oligosaccharides [16]. It is largely known that the substrate specificity and the mode of action of GHs are driven by details of their three dimensional structures rather than by their global fold [17]. Homology modeling is currently routinely used in many practical applications to generate reliable 3D protein structure models from the amino acid sequences. The structural information from the theoretically modelled complex may help us to understand the binding mode between enzymes and substrates leading to an enhanced designing and engineering of novel enzymes with desired structure and functionality for diverse applications.

In our previous study, a gene encoding for a glycoside hydrolase family 1 (GH1) enzyme from the hyperthermophilic *Caldivirga maquilingsensis* strain IC-167 was successfully cloned and expressed in *Escherichia coli* then purified to homogeneity. Its biochemical characterization revealed good substrate specificities towards lactose and other synthetic substrates as well as being significantly thermostable which suggested its potential for oligosaccharides synthesis [18]. Currently, only 13 crystal X-ray structures of glycoside hydrolase β -galactosidases have been elucidated and included in the Protein Data Bank, of which only one belongs to GH 1 family [5]. Therefore in the current study we report on the homology molecular modelling and docking of the substrate into the active site of the GH1 enzyme from the hyperthermophilic *C. maquilingsensis* to understand the enzyme-substrate interaction in relation to GOS synthesis, by using various homologies, docking and verification software.

2. Materials and Methods

2.1. Protein Homology Modelling

The amino acid sequence for the putative gene from *C. maquilingsensis* IC-167 which encoded for a hypothetical

GH1 enzyme has already been deposited in the GenBank (NCBI accession number: ABW01253.1). This amino acid sequence was submitted to Swiss-model web server to predict a three dimensional (3D) structure of the GH 1 from *C. maquilingsensis* against any available template protein structure with the highest amino acid sequence similarity to that of GH 1 from *C. maquilingsensis* in the protein data bank (PDB). The options for the submitted sequence for model construction were selected to be either monomer or tetramer enzyme structure. The generated 3D model structure after searching and comparison with an appropriate template with highest protein structure similarity to the GH1 in the Swiss model server was retrieved from the PDB.

2.2. Three Dimensional Structure Verification

The modeled structure revealed from Swiss model was verified through various verification methods. The structure was checked through PROCHEK to calculate the favored regions, additional allowed regions, generously allowed regions and disallowed regions to predict the protein fold quality. Furthermore, the structure was verified using Errat, to calculate the error values based on the statistics of non-bonded atom-atom interactions in the structure in comparison to a database of reliable high-resolution structures. In addition the structure was checked through Verify 3d to determine the compatibility of an atomic model (3D) with its own amino acid sequence by assigning a structural class based on its location and environment (alpha, beta, loop, polar, nonpolar etc) and comparing the results to good structures.

2.3. Substrate Docking Analysis

The 3D monomer structure generated from Swiss model was then used in the substrate ligand (lactose) docking studies. The docking simulation studies were performed by AutoDock Vina version 5.6. The 3D monomer structure of GH 1 from *C. maquilingsensis* and ligand structure were at first processed using AutoDock software tool with the docking space confined in a grid box centered in the active site of the *C. maquilingsensis* GH1 enzyme. Then a PDBQT file was created for vina software. The low energy pose for best ligand-protein docked model was selected according to results of Vina. The generated output binding mode of the ligand-protein complex was analyzed by Accelrys Discovery studio and Chimera programs to understand and observe the interaction of the ligand and enzyme active site amino acid molecules at an atomic level.

3. Results and Discussion

Three-dimensional structural studies are essential to reveal the identity of important amino acid residues involved in catalysis and to elucidate the mechanism in the enzymatic synthesis of oligosaccharides. It is largely known that the substrate specificity and the mode of action of GHs are driven by details of their three dimensional structures rather than by their global fold [17]. Regarding microbial β -galactosidases, only 13 crystal

structures have been elucidated to date and included in the Protein Data Bank [5]. The structures achieved by homology modeling are said to be similar to those of nuclear magnetic resonance (NMR) spectroscopy or X-ray crystallography [15,19].

When the GH 1 from *C. maquilingensis* amino acid sequence was submitted to the swiss model server to search for a template protein structure with the highest similarity to the target, the most similar template structure was a β -glucosidase from *Sulfolobus Solfataricus* (PDB: 5ixe.4), with 74.69% identity as shown on Figure 1. This result is in agreement with our previous study where we reported GH 1 from *C. maquilingensis* to have 100% β -galactosidase, 50% β -glucosidase and 25% xylosidase enzyme activities and showed a 73% amino acid identity similarity to that of *S. solfataricus*, which implied that these two microorganisms are evolutionary related [18].

Glycoside hydrolases (GH) may be classified in

families based on sequence similarities. Since there is a relationship between folding and sequence similarities, glycoside hydrolases from the same family have similar spatial structures [9,20]. Therefore, this classification facilitates the identification of amino acid residues presenting homologous function in glycoside hydrolases from the same family. Glycoside hydrolases with β -galactosidase (EC 3.2.1.23) activity occur in a variety of microorganisms from the super kingdoms bacteria, archaea and eukaryote and have been categorized within the GH families 1, 2, 35 and 42. GH families 1, 2 and 35 contain enzymes other than β -galactosidases, while only β -galactosidase has been classified into GH 42 [14].

The predicted 3 D model for GH 1 from *C. maquilingensis* was a homo-tetramer with four identical monomer subunits forming a symmetric unit as shown on Figure 3. The overall quality of the predicted 3D model was assessed with PROCHECK, Verify 3d and Errat service tools as shown on Figure 2.

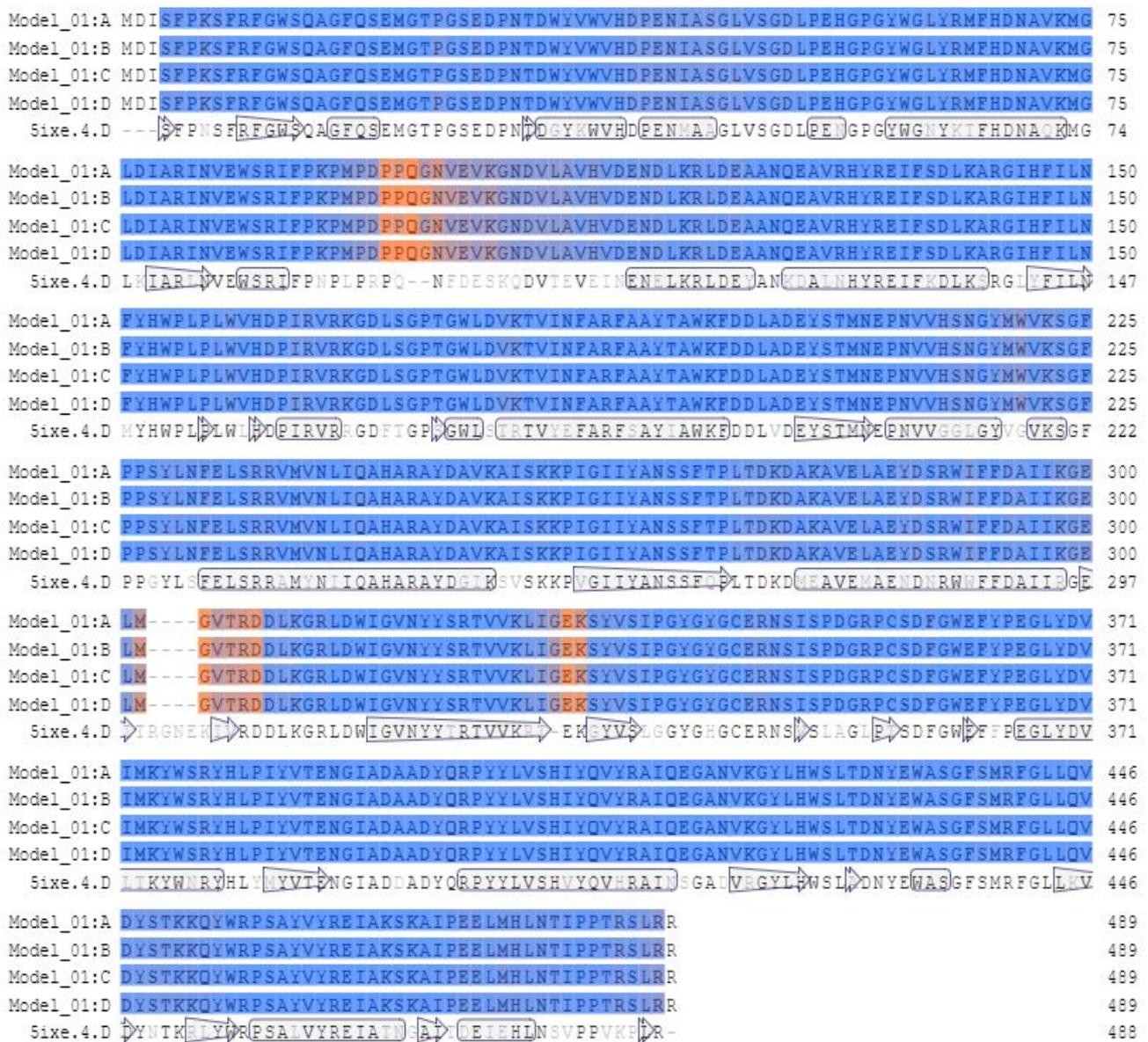


Figure 1. Alignment of amino acid sequences of GH 1 from *C. maquilingensis* and template from β -glucosidase from *Sulfolobus Solfataricus* (PDB: 5ixe.4) through Swiss model

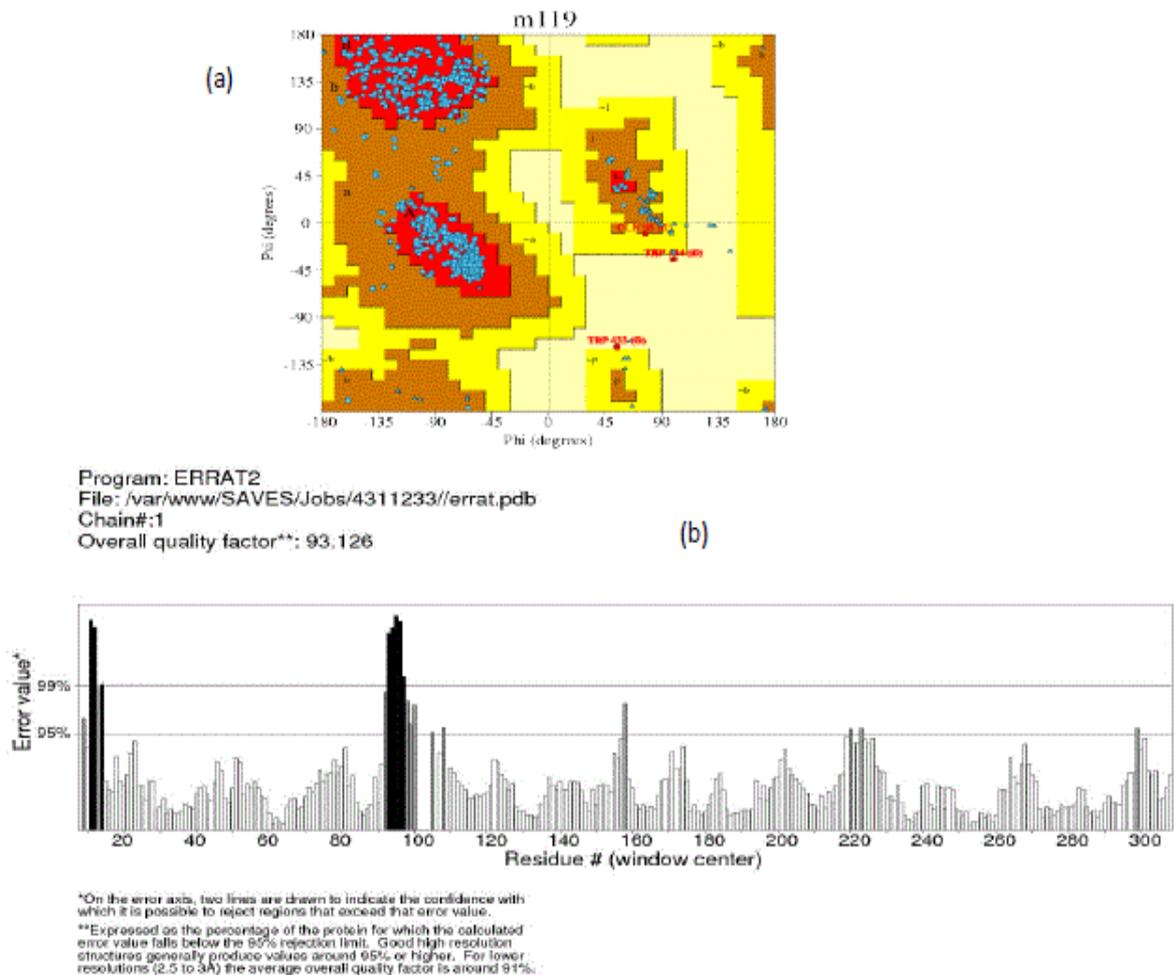


Figure 2. Structure prediction and validation result of GH 1 from *C. maquilingsensis* by homology modeling. (a) Ramachandran plot showing 91.9% of the atoms residing in the most favored region, 7.5% in additional allowed region, 0.1% in the generously allowed region and 0.5% was in the disallowed region. (b) Overall quality of the structure using Errat was 93.126%

The PROCHECK Ramachandran plot statistics for the GH 1 from *C. maquilingsensis* amino acid residues structure atoms revealed that 91.9%, 7.5%, 0.1% and 0.5% atoms resided in the most favored regions, additional allowed regions, generously allowed regions and disallowed regions respectively (Figure 2 (a)). The combined percentage of 99.9 % for all atoms residing in the most favored and allowed regions indicate that there was a good protein fold for the GH 1 from *C. maquilingsensis* structure. When doing the molecular modelling of a crystal x-ray structure of an acidic β -galactosidase from *Aspergillus oryzae*, Maksimainen et al. [20] reported that validation of their model showed 87.8% of the residues in the most favored regions, 11.5% in additional allowed regions, 0.4% in generously allowed regions, and 0.3% in disallowed regions of the Ramachandran plot, which was similar to the results of our model for the GH 1 from *C. maquilingsensis*. The results of PROCHECK for fold validation of our structure further indicated that a high quality GH 1 from *C. maquilingsensis* 3D model structure was successfully derived from *Sulfolobus Solfataricus* (PDB: 5ixe.4) β -glucosidase template in terms of protein folding. The overall quality of the GH 1 from *C. maquilingsensis* structure as displayed by Errat was 93.126% as shown on Figure 2 (b). Further validation of the structure by Verify 3d revealed that 96.70% of the residues had an averaged

3D-1D score ≥ 0.2 for GH 1 from *C. maquilingsensis*.

As already mentioned above, the predicted 3 D model for GH 1 from *C. maquilingsensis* was a homo-tetramer with four identical monomer subunits with half of them forming an asymmetric unit as shown on Figure 3 (a).

The enzyme monomer subunit molecular mass was previously determined by SDS-PAGE at 56.4 kDa [18]. The crystallographic dyad in the tetramer relates the yellow and cyan colored monomers to the blue and red monomers such that it lies flattened perpendicular to one of the 2-fold axes so that each monomer lies at the corner of a slightly puckered square and only contacts two other monomers in the tetramer. The structure forms a typical $(\beta\alpha)_8$ barrel fold with the β -strands and α -helices in each repeat but with substantial long connecting loops in between them. This kind of fold has been observed before and is usually conserved for the family-1 β -glycoside hydrolases [21]. The active site for the GH 1 from *C. maquilingsensis* occurs at the center of the top face of the barrel C-terminal portion and is surrounded by loops connecting the α -helix to the β -strands (Figure 4 (a)). The active site is connected to the surface by a radial tunnel which ends at the center “pocket” of the monomer (Figure 4 (b)), and probably acts as the binding site for extended oligosaccharide substrates. At the active site center of the monomer, there are two access openings or ‘mouths’ connected to the tunnel catalytic site.

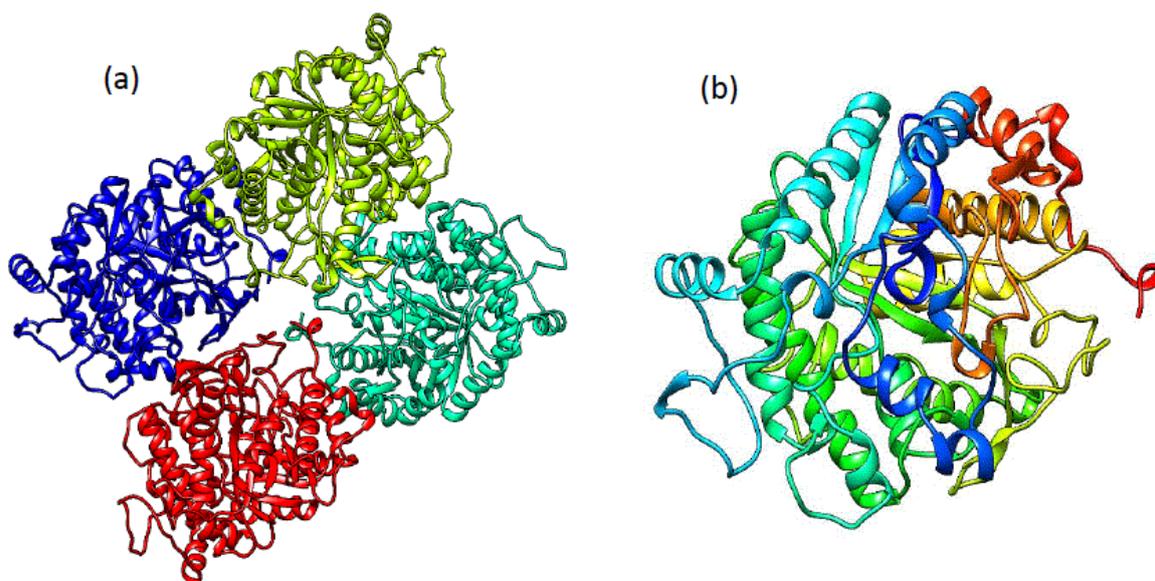


Figure 3. Secondary structure cartoon of the tetramer (a) and the monomer (b) of the GH 1 from *C. maquilingensis*. The crystallographic dyad in the tetramer relates the yellow and cyan monomers to the blue and red monomers forming a perpendicular 2-fold axes

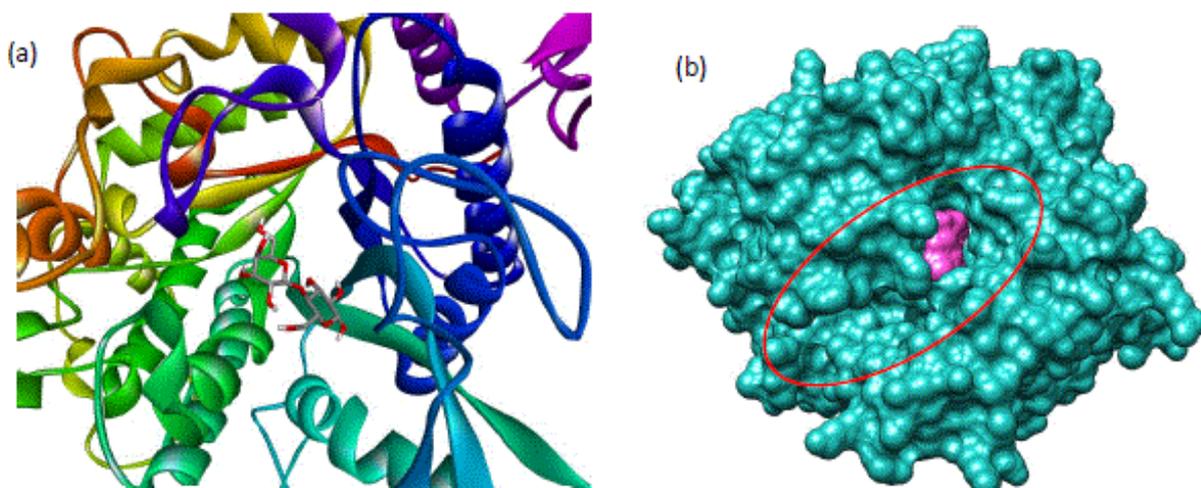


Figure 4. Enzyme and lactose substrate docking at the center catalytic site of GH 1 from *C. maquilingensis* monomer. (a) Cartoon representation showing lactose in sticks representation. (b) Molecular surface representation showing lactose (pink colour) buried in the catalytic pocket (red oval) at the center of monomer, the radial tunnel with two openings on the sides is also visible

Molecular docking can fit two molecules together to obtain an energetically favorable conformation for the complex [22]. The structural information from the theoretically modelled complex are essential to help us to understand and reveal the identity of important amino acid residues involved in catalysis and to elucidate the mechanism in the enzymatic synthesis of oligosaccharides. It has been reported that the substrate specificity and the mode of action of GHs are driven by details of their three dimensional structures rather than by their global fold hence structural analyses and site-directed mutagenesis studies carried out in GHs have revealed that linkage specificity is controlled by the topology of only a few key amino acids located at the enzyme active site [23,24]. Therefore GH 1 from *C. maquilingensis* was docked with lactose substrate as shown on Figure 5.

On the basis of sequence alignment, we identified the active site catalytic amino acid residues in GH 1 from *C.*

maquilingensis as glutamic acid 387 (Glu387), glutamic acid 209 (Glu209) as well as histidine 153 (His 153) with respect to the already determined x-ray crystal structure of *Sulfolobus Solfataricus* [21]. These residues are located in a pocket found at one side of the TIM barrel domain, in the center of each monomer. The glycoside hydrolases active site is divided into several subsites with enough size to bind a monosaccharide unit. As shown in Figure 5, the subsite that binds the galactose moiety of the docked lactose is denominated subsite -1 (or glycone subsite) whereas the remaining part of the substrate, i.e. glucose moiety is accommodated in the aglycone binding region, which may be formed by several subsites (+1, +2, +3 and so on) depending on the degree of polymerization of the acceptor oligosaccharide substrate. Substrate cleavage point is between subsite -1 (glycone subsite) and +1 (aglycone binding region) [9]. The active site contains a high concentration of residues which are highly conserved in all family-1 glycoside hydrolases.

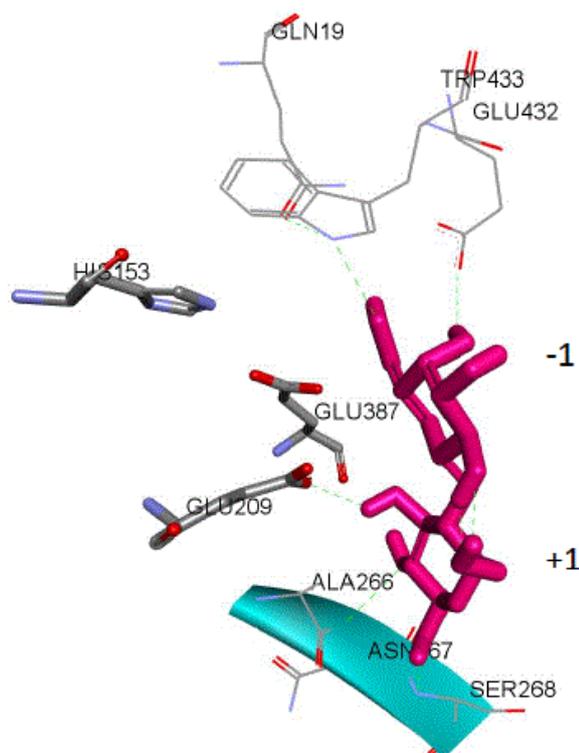


Figure 5. Enzyme/lactose substrate docking solution with lowest free energy value at the catalytic site of GH 1 from *C. maquilingsensis* monomer. Catalytic site residues are shown in stick representation while lactose substrate is colored in pink with subsites -1 and +1. Hydrogen bond interactions with residues surrounding the catalytic site are shown as green dots

In general terms, enzymatic hydrolysis of the glycosidic bond takes place via general acid catalysis that requires two critical residues: a proton donor and a nucleophile/base [17]. As shown on Figure 5, the interactions of lactose with the binding pocket of GH 1 from *C. maquilingsensis* were analyzed using molecular docking tools. The docking of lactose ligand was observed and the interaction with the lowest free energy was chosen as the best orientation position for the enzyme-ligand interaction by the catalytic residues. As shown in the figure, Glu387 and Glu209 are the nucleophile and acid/base catalyst for GH 1 from *C. maquilingsensis*, respectively. The nucleophile is responsible for formation of a covalent galactosyl-enzyme intermediate while the acid/base is a proton donor. Briefly the proposed catalysis proceeds via a double-displacement reaction involving two steps of hydrolysis/glycosylation and transglycosylation. In the first step, the enzyme is glycosylated by the concerted action of the two carboxylate residues, Glu387 and Glu209, in which a covalent galactosyl-enzyme intermediate is formed and hydrolyzed via oxocarbenium ion-like transition states, while glucose is being released. The second mechanistic step involves the attack of the covalent galactosyl-enzyme intermediate by a water molecule (hydrolysis) or a sugar molecule (transgalactosylation), followed by the transfer of a proton from a water/sugar to the proton donor, in a reverse mode of the first step [25]. The attacking from another sugar molecule results in synthesis of galactooligosaccharides by transgalactosylation whereas a water molecule attack results in hydrolysis products glucose and galactose. It is worth noting that

other residues like, alanine (Ala 266), asparagine (Asp267), serine (Ser268), Glutamic acid (Glu432), tryptophan (Trp433) and glutamine (Gln19) found surrounding the catalytic site are within hydrogen bonding distances with the glycone and aglycone playing essential roles in ligand binding and recognition of the lactose molecules and in selecting different acceptor molecules during transgalactosylation unique to this enzyme. In fact the hydroxyl, polar and aromatic side chains found in some of these residues have been reported to provide flexibility and fine-tuning of the position of the nucleophile and also stabilizing the deprotonated state of the substrate-enzyme intermediate which is beneficial for both transgalactosylation and hydrolysis [20,26]. Therefore, the orientation and establishment of bonds between the substrate and both aromatic and polar amino acids define the relevant enzymatic features, such as the affinity, specificity, and activity. There are also a number of negatively charged and polar acidic glutamic acid (Glu) residues around the active site area implying that this enzyme has a good structural basis to act in acidic pH range. Indeed, previously the activity analysis for the GH 1 from *C. maquilingsensis* showed that it had a pH optimum at about 4.5 [18].

4. Conclusion

Molecular docking simulation elucidated that GH 1 from *C. maquilingsensis* has a tetrameric structure with an active site at the center of its monomer with a tunnel that is connected to the catalytic “pocket”. The structural data was used to identify amino acid residues involved in substrate catalysis and having interactions with other residues surrounding the active site.

The findings offer an improved understanding of how enzyme protein structure determines catalytic specificity, which serves as new knowledge basis to engineer GH 1 from *C. maquilingsensis* for the biosynthesis of GOS with a broad or narrow degree of polymerization range. Some of the known glycoside hydrolases are quite specific, whereas others tolerate some slight structural differences. The characterization of the molecular basis of glycoside hydrolases substrate preference may contribute to the comprehension of the enzymatic specificity, a fundamental property of biological systems. Therefore, not only structure but binding site topology are very important factors for understanding the mechanism of recognition of these proteins.

Acknowledgements

This work was supported by the Natural Science Foundation of China Project (No. 31371788 and No. 21276001), National High Technology Research and Development Program of China (2013AA102102).

References

- [1] Al-Sheraji, S. H., Ismail, A., Manap, M. Y., Mustafa, S., Yusof, R. M. & Hassan, F. A. Prebiotics as functional foods: A review, *Journal of Functional Foods*. 5, 1542-1553, 2013.

- [2] Gänzle, M. G. Enzymatic synthesis of galacto-oligosaccharides and other lactose derivatives (hetero-oligosaccharides) from lactose, *International Dairy Journal*. 22, 116-122, 2012.
- [3] Intanon, M., Arreola, S. L., Pham, N. H., Kneifel, W., Haltrich, D. & Nguyen, T. H. Nature and biosynthesis of galacto-oligosaccharides related to oligosaccharides in human breast milk, *FEMS microbiology letters*. 353, 89-97, 2014.
- [4] Sangwan, V., Tomar, S. K., Singh, R. R., Singh, A. K. & Ali, B. Galactooligosaccharides: novel components of designer foods, *Journal of food science*. 76, R103-11, 2011.
- [5] Diez-Municio, M., Herrero, M., Olano, A. & Moreno, F. J. Synthesis of novel bioactive lactose-derived oligosaccharides by microbial glycoside hydrolases, *Microbial biotechnology*. 7, 315-31, 2014.
- [6] Otieno, D. O. Synthesis of β -Galactooligosaccharides from Lactose Using Microbial β -Galactosidases, *Comprehensive Reviews in Food Science and Food Safety*. 9, 471-482, 2010.
- [7] Nath, A., Verasztó, B., Basak, S., Koris, A., Kovács, Z. & Vatai, G. Synthesis of Lactose-Derived Nutraceuticals from Dairy Waste Whey-a Review, *Food and Bioprocess Technology*. 9, 16-48, 2015.
- [8] Rodriguez-Colinas, B., Fernandez-Arrojo, L., Ballesteros, A. O. & Plou, F. J. Galactooligosaccharides formation during enzymatic hydrolysis of lactose: towards a prebiotic-enriched milk, *Food Chem*. 145, 388-394, 2014.
- [9] Marana, S. R. Molecular basis of substrate specificity in family 1 glycoside hydrolases, *IUBMB life*. 58, 63-73, 2006.
- [10] Moracci, M., Cabaldo, L., Ciamarella, M. & Rossi, M. Identification of two glutamic acid residues essential for catalysis in the b-glycosidase from the thermoacidophilic archaeon *Sulfolobus solfataricus*, *Protein Engineering*. 9, 1191-1195, 1996.
- [11] Torres, D. P. M., Gonçalves, M. d. P. F., Teixeira, J. A. & Rodrigues, L. R. Galacto-Oligosaccharides: Production, Properties, Applications, and Significance as Prebiotics, *Comprehensive Reviews in Food Science and Food Safety*. 9, 438-454, 2010.
- [12] Park, A. R. & Oh, D. K. Galacto-oligosaccharide production using microbial beta-galactosidase: current state and perspectives, *Appl Microbiol Biotechnol*. 85, 1279-86, 2010.
- [13] Duarte, P. M., Goncalves, M. P., Teixeira, J. A. & Rodrigues, L. R. Galacto-oligosaccharides: production, properties, applications, and significance as prebiotics, *Compre Rev Food Sci Food Safety*. 9, 438-454, 2010.
- [14] Gosling, A., Stevens, G. W., Barber, A. R., Kentish, S. E. & Gras, S. L. Recent advances refining galactooligosaccharide production from lactose, *Food Chemistry*. 121, 307-318, 2010.
- [15] Sanchez, R., Pieper, U., Melo, F., Eswar, N., Marti-Renom, M. A., Madhusudhan, M. S., Mirkovic, N. & Sali, A. Protein structure modeling for structural genomics, *Nat Struct Biol* 7, 986-990, 2000.
- [16] Henrissat, B. & Davies, G. Structural and sequence-based classification of glycoside hydrolases *Curr Opin Struct Biol*. 7, 637-644, 1997.
- [17] Davies, G. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases *Structure*. 3, 853-859, 1995.
- [18] Letsididi, R., Hassanin, H. A., Koko, M. Y., Ndayishimiye, J. B., Zhang, T., Jiang, B., Stressler, T., Fischer, L. & Mu, W. (2016) Characterization of a thermostable glycoside hydrolase (CMbg0408) from the hyperthermophilic archaeon *Caldivirga maquilungensis* IC-167, *J Sci Food Agric*. 97: 2132-2140, 2017.
- [19] Sánchez, R. & Šali, A. "Advances in comparative protein-structure modelling" *Current opinion in structural biology*. 7, 206-214, 1997.
- [20] Maksimainen, M. M., Lampio, A., Mertanen, M., Turunen, O. & Rouvinen, J. The crystal structure of acidic beta-galactosidase from *Aspergillus oryzae*, *International journal of biological macromolecules*. 60, 109-15, 2013.
- [21] Aguilar, C. F., Sanderson, I., Moracci, M., Ciamarella, M., Nucci, R., Rossi, M. & Pearl, L. H. Crystal Structure of the b-Glycosidase from the Hyperthermophilic Archeon *Sulfolobus solfataricus*: Resilience as a Key Factor in Thermostability, *J Mol Biol* 271, 789-802, 1997.
- [22] Meng, X. Y., Zhang, H. X., Mezei, M. & Cui, M. Molecular docking: A powerful approach for structure-based drug discovery *Current Computer-Aided Drug Design*. 7, 146-157, 2011.
- [23] Yang, J., Wang, Q., Zhou, Y., Li, J., Gao, R. & Guo, Z. Engineering *T. naphthophila* β -glucosidase for enhanced synthesis of galactooligosaccharides by site-directed mutagenesis, *Biochemical Engineering Journal*. 127, 1-8, 2017.
- [24] Wu, Y., Yuan, S., Chen, S., Wu, D., Chen, J. & Wu, J. Enhancing the production of galacto-oligosaccharides by mutagenesis of *Sulfolobus solfataricus* beta-galactosidase, *Food Chem*. 138, 1588-95, 2013.
- [25] Brás, N. F., Fernandes, P. A. & Ramos, M. J. QM/MM studies on the β -galactosidase catalytic mechanism: hydrolysis and transglycosylation reactions, *J Chem Theory Comput* 6, 421-433, 2010.
- [26] Armenta, S., Moreno-Mendieta, S., Sanchez-Cuapio, Z., Sanchez, S. & Rodriguez-Sanoja, R. Advances in molecular engineering of carbohydrate-binding modules, *Proteins*. 85, 1602-1617, 2017.