Research Article

Journal of Proteins and Proteomics

Structural characterization of an endo β-1,3-glucanase of family 81 glycoside hydrolase (*Ct*Lam81A) from *Clostridium thermocellum*

Krishan Kumar, Priyanka Nath and Arun Goyal[*]

Carbohydrate Enzyme Biotechnology Laboratory, Department of Bioscience and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India

***Corresponding author:** Arun Goyal. Tel: +91-361-258-2208
E-mail: arungoyl@iitg.ernet.in.

Running title: Structural characterization of *Ct*Lam81A

Received: February 13, 2018
Accepted: August 15, 2018
Published: August 20, 2018

Abstract

The gene sequence of a family 81 glycoside hydrolase from *Clostridium thermocellum Cthe*_0660, (GenBank accession number: ABN51485.1) was retrieved from CAZy (Carbohydrate-Active Enzyme) database. The gene (2151 bp) encoding the catalytic module of *Cthe*_0660, an endo β-1,3-glucanase (*Ct*Lam81A) cloned, expressed earlier was used in this study. The secondary structure analysis of *Ct*Lam81A by PsiPred showed the presence of 23% α-helices and 23% β-strands whereas, CD analysis showed 24% of α-helices and 24% of β-strands. The secondary structure analysis by Circular Dichroism corroborated the PsiPred results. The 3D modeled structure of *Ct*Lam81 generated by Swiss-Model server and energy minimized by Swiss-PDB viewer displayed an N-terminal β-sandwich domain, a $(\alpha/\alpha)_6$ domain and a short β-sandwich domain at C-terminal. The quality assessment of modelled structure by Ramachandran plot displayed 94% residues in favoured region, 4.6% in allowed region and 1.4% in outlier region. The superposition of *Ct*Lam81 modeled structure with its closest homologue *Bh*0236 (PDB ID: 5t49) from *Bacillus halodurans* indicated that Glu515 acts as a catalytic base, while Asp438 acts as a catalytic acid. Multiple sequence alignment showed that these catalytic residues are conserved within the family 81 glycoside hydrolase. *Ct*Lam81showed lowest binding energy against laminaripentaose and the amino acid residues, Tyr363, Trp364, Tyr430, Trp588, Trp589 and Trp637create the binding pocket at active site to hold the ligand.

Key words: Family 81 glycoside hydrolase; β-1,3-glucanase; Laminari-oligosaccharides; Circular Dichroism; Homology modeling; Docking; *Clostridium thermocellum*

Note - Coloured Figures are available on the Journal Website in "Archives" Section

Introduction

β-1,3-glucan is a major carbohydrate found in bacteria (*Alcaligenesfaecalis, Agrobacteriumrhizogenes, Agrobacterium radiobacter*), as a linear chain of β-1,3-linked glucose (McIntosh et al., 2005) while in fungi (*Lentinusedode, Sclerotiumrolfsii, schizophyllum commune*) and algae (*Laminariadigitata)*a chain ofβ-1,3-glucan branched either with β-1,6-linked glucose or with short chain of β-1,3-glucan (McIntosh et al., 2005; Volmanet al., 2008). β-1,3-glucan is hydrolysed by β-1,3-glucanase enzyme, which are classified in to two types, exo- and endo- acting. Exo-β-1,3-glucanase (EC 3.2.1.58), acts on the terminal of chain and releases glucose ordimers (Mouynaet al., 2013). These exo-acting enzymes fall into the glycoside hydrolase families 3, 5, 17 and 55. Endo-β-1,3-glucanase (EC 3.2.1.39) acts randomly on the β-1,3-glucans and releases mixture of oligosaccharides. Endo-β-1,3-glucanases falls in to families 16, 17, 55, 64, 81 and 128 of glycoside hydrolase (GH). Thus, both GH17 and GH55 families contain β-1,3-glucan specific enzymes that have both, exo- or endo-activities. Family 81 glycoside hydrolases are known for their endo-β-1,3-glucanase activity and are widely distributed in plants, yeast, fungi, bacteria, archaea and viruses. Presently, family 81 GH contains 674 proteins distributed over bacteria, archaea, eukaryotes and viruses (http://www.cazy.org/GH81.html). All of the proteins in family 81 GHs share a common size of around 650 amino acids. Some of its members are biochemically characterized like β-glucan-binding protein (GBP) from Glycine max (Fliegmann et al., 2005), *Tf*Lam81 from *Thermobifida fusca* (McGrath & Wilson, 2006) and Eng2 from *Saccharomyces cerevisiae* (Martin-Cuadrado et al., 2008). ^{1}H NMR spectroscopic analysis of GBP hydrolysed product revealed the retaining hydrolytic mechanism of this family (Fliegmannet al., 2005). So far only two enzymes, *Rm*Lam81A from *Rhizomucor miehei* (Zhou et al., 2013) from *Bacillus halodurans* (Pluvinageet al., 2017) provide information about crystal structure solved from family GH81. Crystal Structure analysis of

*Rm*Lam81A depicted the presence of β-sandwich domain at N-terminal and $(\alpha/\alpha)_6$ domain at C-terminal connected *via* additional domain between them. Similar type of domain was also found in *Bh*GH81 protein. Interaction study of laminarin with *Bh*GH81 revealed that the active site of this protein can accommodate maximum five glucose residues of β-1,3-glucan (Pluvinage et al., 2017). *Clostridium thermocellum* contains multienzyme complex, known as cellulosome. Cellulosome complex is made by a high affinity interaction between dockerin modules of the cellulosomal catalytic proteins and cohesin modules on the scaffoldins (Fontes and Gilbert, 2010). One such cellulosomal enzyme of *Clostridium thermocellum*is β-1,3-glucanase (*Cthe*_0660) belongs to family 81 GH. The catalytic component, *Ct*Lam81A of *Cthe*_0660 was cloned and purified earlier (unpublished results) was used in the present study. In this study, the secondary structure elements of *Ct*Lam81A were determined by Circular Dichroism and also confirmed by PsiPred. The 3D structure of *Ct*Lam81A was modeled by homology modeling and ligand docking study was performed to identify the amino acid residues involved in the catalysis.

Material and Methods

Retrieval of amino acids sequence of CtLam81A - The gene sequence of family81 glycoside hydrolase protein of *Clostridium thermocellum* ATCC27405 having gene accession number ABN51895.1 and uniport ID A3DD6 was retrieved from CAZy database (http://www.cazy.org/). The boundaries of the family 81 GH catalytic domain (*Ct*Lam81A) in the amino acid sequence of Cthe_0660 protein was identified by using the conserved domain database (http://www.ncbi.nlm.nih.gov/cdd/). The 2151 bp gene encoding *Ct*Lam81A protein was cloned, expressed and purified earlier in our laboratory (Unpublished results).

Secondary structure analysis of CtLam81A - Secondary structure of *Ct*Lam81A protein sequence for alpha helices, beta strands, turns and loops were predicted with the help of

PsiPred v3.3 method (http://bioinf.cs.ucl.ac.uk/psipred/). The composition of secondary structure of *Ct*Lam81A was also determined by Circular dichroism (CD). The purified *Ct*Lam81 at a concentration 0.7μM in 50mM phosphate buffer, pH 7.5 was used for CD analysis. The CD spectrum was recorded on a spectro-polarimeter (JascoJ-815, Japan) at 25°C using 1 nm bandwidth over far UV region between 190 to 250 nm at scanning rate of 50 nm/min. The CD data were presented by difference in molar extinction coefficient (Δε, deciliter mol^{-1} cm^{-1}) as a function of wavelength (Kelly et al., 2005). Percentage of α-helix and β-sheet were measured by web based K2D3 software package (http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/) (Andrade et al., 1993).

Homology modelling CtLam81A-The three dimensional structure of *Ct*Lam81A was modelled with the help of SWISS-MODEL server. SWISS-MODEL is a web based server, which automatically identify structural templates based on BLAST analysis and generate 3-D protein structure (https://swissmodel.expasy.org/). The best modelled structure was selected based on the alignment score with templates.

Model refinement and quality assessment - The best modeled structure was energy minimized by Swiss-PDB viewer. Swiss-PDB viewer comprises a version of GROMOS 43B1 force field (Van Gunsteren et al., 1996). This force field allows the evaluation of the energy of structure as well as it repairs the distorted geometry through energy minimization. Then the quality of energy minimized structure was checked by plotting the Ramachandran plot using RAMPAGE server (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php). This modeled structure was also validated by using the Verify-3D program (http://services.mbi.ucla.edu/Verify_3D/), which analyzes the compatibility of atomic model with the amino acid sequence (Liithy et al., 1992).

Prediction of active site and mechanism of action - The amino acid residues in the active site are generally conserved in the proteins of the same family. To identify the amino acid

residues at active sites, *CtLam*81A was structurally aligned with the structure of *Bh*GH81 (PDB ID- 5T49) using Pymol. To confirm whether these residues are conserved within family 81 GHs, multiple sequence alignment (MSA) was performed with the already characterized homologous proteins of the family81 GHs *viz*. *Bacillus halodurans* GH81 (*Bh*GH81; Q9KG76), *Thermobifida fusca* Lam81A (*Tf*Lam81A, AAZ56163), *Glycine max* Gbp (*Gm*Gbp; BAA11407), *Aspergillus fumigatus* EngA (*Af*EngA; AAF13033), *Saccharomyces cerevisiae* Eng2 (*Sc*Eng2; AAB82378) and *Rhizomucor miehei* Lam81 (*Rm*Lam81A; KC847083). The percentage similarity between amino acid sequences of *Ct*Lam81A with its homologous proteins from family 81 GH was analysed by BLAST tool (Altschul et al., 1990). MSA was performed with the help of CLUSTALW program (http://www.genome.jp/tools-bin/clustalw) and the conserved amino acid residues were viewed by ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/). To study the mechanism of hydrolysis the distance between the two catalytic amino acids (Asp438 and Glu515; from C_{δ}toC_{δ}) of *Ct*Lam81A was measured in the molecular visualization system PyMol 2.0.

Molecular dynamic simulation of CtLam81A modeled structure - Molecular dynamic (MD) simulation of *Ct*Lam81A was performed by using GROMACS v 5.14 (Berendsen et al, 1995). GROMOS96 53a6 force field was used to calculate the Protein forces, where the protein *Ct*Lam81A was placed within in a cubic box of single point charge (SPC) with water molecules. The *Ct*Lam81A protein charges were neutralized by adding the 15 numbers of counter ions (Na^{+}). Then the whole system was equilibrated for 500 ps in NVT ensemble (constant number of particles, volume and temperature) was carried out for restraining the solute atoms. This system was again equilibrated for 500 ps by NPT ensemble (constant number of particles, pressure and temperature) twice, once with restraints and then without restraints. Production run was performed for 30 ns with NPT ensemble adopting a 2 fs integration time. The linear constraint solver (LINCS) algorithm (Hess et al, 1997) was

employed to constrain the bonds associated with hydrogen atoms and radius of gyration. Throughout the production run the modeled *Ct*Lam81A structure was analyzed as a time dependent function to ascertain its stability in the solvent system. The variation in the *Ct*Lam81A protein backbone (root mean square deviation, RMSD) was estimated by the least square fitting method.

Biding interaction analysis of CtLam81A with the ligands - Molecular docking of *Ct*Lam81A with the laminari oligosaccharides was performed by using SwisDock, the web based server (http://www.swissdock.ch/docking). Laminaribiose, laminaritriose and laminaritetraose were downloaded from PubChem database (https://pubchem.ncbi.nlm.nih.gov/). Laminaripentaose and laminarihexaose were generated using the GLYCAM server (Kirschneret al., 2008). The modelled *Ct*Lam81A was saved in PDB format and the ligands were saved in Mol2 file format for docking analysis in the Swiss Dock tool. Swiss Dock generates large number of ligand binding results. The enzyme-ligand docked complex showing the strongest binding with maximum of negative binding energy were selected. This ligand bound structure was downloaded and visualized in PyMol 2.0. The depiction of ligand interaction with the amino acid residues of the protein was generated using the PDBsum Generate tool (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate.html).

Results and Discussion

Molecular architecture of CtLam81A

*Ct*Lam81A from *Clostridium thermocellum* contains an N-terminal catalytic domain and a type I dockerin domain at the C-terminal connected via linker. The dockerin domain binds to cohesion domain, located on the scaffoldin protein in cellulosome complex (Ding et al, 2001). The 2151 bp gene encoding family 81 glycoside hydrolase (*Ct*Lam81) and linker was cloned in pET-28a(+) vector, expressed in *E. coli*BL21 cells and purified (Unpublished

results). Total length of purified protein(*Ct*Lam81A) is 740 amino acids consisting of 23 amino acids of pET28a(+) vector, 660 amino acids of catalytic domain and 64 amino acids of linker at the C-terminal. The theoretical molecular size of *Ct*Lam81A is 82.4 kDa.

Secondary structure analysis of CtLam81A

Secondary structure of *Ct*Lam81A predicted by PsiPred tool showed the presence of 14 α-helices (23%), 31 β-strands (23%) and 45 random coils (54%) (Fig.1). *Ct*Lam81A showed almost an equal amount of β-strands and α-helices. The N-terminal of *Ct*Lam81A showed the predominance of β-strands and the C-terminal showed the abundance of α-helices. The predicted secondary structure was also confirmed by circular dichroism analysis of purified *Ct*Lam81A. The CD spectrum of *Ct*Lam81A was analysed with the help of K2D3 server, which gives the percentage of α helix and β strand by comparing the available secondary structures of known proteins (Louis-Jeuneet al., 2012). This analysis showed that *Ct*Lam81A contains 24% α-helices and 24% β-strands (Fig. 2). The results of CD analysis corroborated with the secondary structure predicted by PsiPred method. This distribution of secondary structure elements is common in family 81 GH (Zhou et al., 2013).

Homology modeling and structure validation

The web based server SWISS-MODEL generated the 3-D model structure of *Ct*Lam81A by using the structural template of protein *Bh*GH81(PDB ID- 5T49) having 42% sequence identity with query coverage 100%. The overall structure of the modelled *Ct*Lam81A (Fig. 3A) consisted of three domains, N- terminal β-sandwich domain I (red colour), an $(\alpha/\alpha)_6$barrel domain II(green colour) and a shortβ-sandwich domain III (yellow colour) at C-terminal. This modeled structure was energy minimized and validated before docking analysis. Ramachandran plot of *Ct*Lam81A modeled protein showed that 94% of the residues found in the favourable region and 4.6 % of residues found in allowed region and only 1.4% residues (Phe177, Thr189, Lys221, Asn465, Asp481, Thr592, Ile661 and Pro663)

was found in outlier region (Fig. 3B). This indicated that the amino acid residues in the modelled *Ct*Lam81A occupied favourable phi (φ) and psi (ψ) backbone dihedral angles. Verify_3D result showed that 95.25% of the residues in modeled structure had an averaged 3D-1D score≥ 0.2, which indicated the compatibility of amino acids with the modelled structure (Fig. 3C). ProSA result of modelled *Ct*Lam81A indicated that the protein is error free and reside in the x-ray zone with Z-scores of −9.88 (Fig. 3D).

Catalytic cleft and mechanism of catalysis ofCtLam81A

The amino acids residues involved in the catalysis and in substrate binding were explored by superimposing the *Ct*Lam81A modelled structure with crystal structure of*Bh*GH81 (PDB ID – 5T49) having RMSD, 0.495 Å as shown in Fig. 4A. The catalytic core of *Ct*Lam81A existed within $(\alpha/\alpha)_6$ barrel domain of the protein. The superposition of modelled structure of *Ct*Lam81A (shown in green colour) with *Bh*GH81 (shown in cyan colour) showed that β-sheet and α-helix of *Ct*Lam81A aligned well with the β-sheet and α-helix of *Bh*GH81 (Fig. 4A). However, some loop regions (43-65 amino acids, 96-105 amino acids, 214-222 amino acids and 296–302 amino acids) of the *Ct*Lam81A modelled structure do not align with the *Bh*GH81 structure. The analysis of superimposition showed that Glu515 acts as catalytic base and Asp438 acts as acid catalyst while Trp588 and Trp589 helps in binding the ligand (Fig. 4B). *Ct*Lam81A shared amino acid sequence similarity with homologous proteins from *Bacillus halodurans* GH81 (*Bh*GH81; Q9KG76), *Thermobifida fusca* Lam81A (*Tf*Lam81A, AAZ56163), *Glycine max* Gbp (*Gm*Gbp; BAA11407), *Aspergillus fumigatus* EngA (*Af*EngA; AAF13033), *Saccharomyces cerevisiae* Eng2 (*Sc*Eng2; AAB82378) and *Rhizomucor miehei* Lam81 (*Rm*Lam81A; KC847083) as shown in Table 1. The multiple sequence alignment of *Ct*Lam81A showed that the catalytic amino acids are conserved in all family 81 glycoside hydrolases. Trp588is conserved in both bacterial and plant proteins, while Trp589 is conserved in bacteria only (Fig. 5). Glycoside

hydrolases catalyse the glycosidic bond, either by inverting or retaining hydrolytic mechanism. It has been found that, in the inverting mechanism, the average distance between two carboxyl groups of acid/base catalytic residues is 10.5 Å and in the retaining mechanism the average distance is only 5.5 Å (Zechel& Withers, 1999). The superimposition of *Ct*Lam81A structure with crystal structure of *Bh*GH81 indicated that Glu515 is catalytic base and Asp438 is catalytic acid. The distance between these two catalytic residues was found to be 8.3Å (Fig. 6), indicating the inverting hydrolytic mechanism for *Ct*Lam81A.The same inverting hydrolytic mechanism was also reported for β-glucan-binding protein (GBP) a member of family 81 glycoside hydrolases from soybean (Fliegmanet al., 2005).

Molecular dynamics simulation of CtLam81A modeled structure

Molecular dynamic simulation of modeled *Ct*Lam81A was performed in order to analyse the structure stability and compactness of the structure over a 30 ns duration. The MD simulation results showed the fluctuation in RMSD value of *Ct*Lam81A up to 5 ns and thereafter the structure was completely stable till 30 ns. The overall deflection was 0.32 nm RMSD (Fig. 7A). The radius of gyration (Rg) of *Ct*Lam81A structure remained between 2.31 and 2.39 nm till the end of the run (Fig. 7B). Based on the MD simulation result of CtLam81A, it was concluded that the modelled structure of *Ct*Lam81A structure has a stable conformation.

Docking study of CtLam81A with ligands

The docking study of *Ct*Lam81A was performed in order to analyse the interaction of ligands with its active site. The binding energies of laminari-oligosaccharides with *Ct*Lam81A are shown in Table 2. Laminaripentaose showed the strongest binding with maximum of negative binding energy of -11.47 kcal/mol. The docking of laminaripentaose on surface of *Ct*Lam81A is shown in Fig. 8A. The non-reducing end of laminaripentaose is expressed as *–n* subsite and the reducing end as *+n* subsite. This result revealed that the

catalytic cavity of *Ct*Lam81A can occupy maximum of five glucose residues of β-1,3 glucan. Similar results were reported for β-1,3-glucanase (*Bh*GH81) from *Bacillus halodurans* (Pluvinageet al., 2017). The aromatic amino acid residues *viz.* Tyr363, Trp364, Tyr430, Trp588, Trp589 and Trp637 in the catalytic pocket are involved in the holding of laminaripentaose. Trp 589 of *Ct*Lam81A lies near the -2 subsite of the laminaripentaose and Trp 588 lies near the +1 subsite, which helps in holding the ligand, as also reported for *Bh*GH81 from *Bacillus halodurans* (Pluvinageet al., 2017). The catalytic core residues, Tyr363, Lys367, Tyr430, Asp438, Asp513 and Glu519 of *Ct*Lam81A are making polar interactions with laminaripentaose (Fig. 8B), while, the residues Trp364, Gly431, Phe441, Ala502, Glu515, Trp588, Ile599, Leu602, Trp637 and Asp665 present at the catalytic core are involved in hydrophobic interactions with the ligand.

Conclusion

The amino acid sequence of *Ct*Lam81A was analysed for secondary structure elements by PsiPred tool and CD. CD analysis showed 24% of α-helices and 24% of β-strands. Modelled structure of *Ct*Lam81A showed the presence of an N-terminal β-sandwich domain, a $(\alpha/\alpha)_6$ domain and a short β-sandwich domain at C-terminal. The active site lies in the $(\alpha/\alpha)_6$ domain of the protein. Superimposition of *Ct*Lam81A with *Bh*GH81 and MSA confirmed that Glu515 is catalytic base and Asp438 is catalytic acid, which are 8.3 Å apart and involved in hydrolysing β-1,3-glucan by inverting mechanism. The docking study of *Ct*Lam81A with ligands showed the strongest binding energy against laminaripentaose, because the active site pocket of *Ct*Lam81A can occupy 5 glucose residues of β-1,3-glucan. The aromatic amino acid residues Tyr363, Trp364, Tyr430, Trp588, Trp589 and Trp637, Trp 589 and Trp 588 create a binding pocket for holding the ligand at the catalytic site.

Acknowledgements

This research work was carried out under bilateral joint DST (Ministry of Science and Technology, India) project (No. DST/INT/Portugal/P-14/2013) in collaboration with FCT Lisbon, Portugal. The authors thank Prof. T. Punniyamurthy, Head, Department of Chemistry, Indian Institute of Technology Guwahati for providing CD facility.

Abbreviations

GH, glycoside hydrolase; CD, Circular Dichroism

Conflict of Interest

The authors declare no conflict of interest.

References

Altschul, S.F., Gish, W., Miller. W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol. *215,* 403-410.

Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993). Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Protein Eng Des Sel. *6,* 383-390.

Berendsen, H. J., van der Spoel, D., & van Drunen, R. (1995). GROMACS: a message-passing parallel molecular dynamics implementation. Comput Phys Commun. *91,* (1-3), 43-56.

Ding, S. Y., Rincon, M. T., Lamed, R., Martin, J. C., McCrae, S. I., Aurilia, V., and Flint, H. J. (2001). Cellulosomal Scaffoldin-Like Proteins from *Ruminococcus flavefaciens*. J Bacteriol., *183,* 1945-1953.

Fliegmann, J., Montel, E., Djulić, A., Cottaz, S., Driguez, H. andEbel, J. (2005) Catalytic properties of the bifunctional soybean β-glucan-binding protein, a member of family 81 glycoside hydrolases. FEBS Letters. *579*, 6647-6652.

Fontes, C. M., and Gilbert, H. J. (2010). Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. Annu Rev Biochem. *79*, 655-681.

Hess, B., Bekker, H., Berendsen, H. J., & Fraaije, J. G. (1997). LINCS: a linear constraint solver for molecular simulations. J comp chem. *18*, 1463-1472.

Kelly, S. M., Jess, T. J., and Price, N. C. (2005). How to study proteins by circular dichroism. BBA Proteins and Proteomics. *1751*, 119-139.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem *29*, 622-655.

Liithy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. Nature. *356*, 83-85.

Louis-Jeune, C., Andrade-Navarro, M. A., and Perez-Iratxeta, C. (2012). Prediction of protein secondary structure from circular dichroism using theoretically derived spectra. Proteins: Struct Funct Bioinf. *80*, 374-381.

Martín-Cuadrado, A. B., Fontaine, T., Esteban, P. F., del Dedo, J. E., de Medina-Redondo, M., del Rey, F. and de Aldana, C. R. V. (2008) Characterization of the endo-β-1, 3-glucanase activity of *S. cerevisiae* Eng2 and other members of the GH81 family. Fungal Genet Biol. *45*, 542-553.

McGrath, C. E. and Wilson, D. B. (2006) Characterization of a *Thermobifidafusca*β-1, 3-glucanase (Lam81A) with a potential role in plant biomass degradation. Biochemistry. *45*, 14094-14100.

McIntosh, M., Stone, B. A. and Stanisich, V. A. (2005) Curdlan and other bacterial (1→ 3)-β-D-glucans. Appl Microbiol Biotechnol. *68*, 163-173.

Mouyna, I., Hartl, L. and Latgé, J. P. (2013) β-1,3-glucan modifying enzymes in *Aspergillus fumigatus*. Front Microbiol. *4*, 81.

Pluvinage, B., Fillo, A., Massel, P., and Boraston, A. B. (2017). Structural analysis of a family 81 glycoside hydrolase implicates its recognition of β-1,3-glucan quaternary structure. Structure. 25, 1348-1359.

Van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hünenberger, P. H., Krüger, P. K. H. C., Mark, A. E., and Tironi, I. G. (1996). Biomolecular simulation: the {GROMOS96} manual and user guide. VdFHochschulverlag ETHZ, Zurich.

Volman, J. J., Ramakers, J. D. and Plat, J. (2008) Dietary modulation of immune function by β-glucans. Physiol Behav. *94*, 276-284.

Zechel, D. L., and Withers, S. G. (2000). Glycosidase mechanisms: anatomy of a finely tuned catalyst. Acc Chem Res. *33*, 11-18.

Zhou, P., Chen, Z., Yan, Q., Yang, S., Hilgenfeld, R., and Jiang, Z. (2013). The structure of a glycoside hydrolase family 81 endo-β-1, 3-glucanase. Acta Crystallogr D Biol Crystallogr. *69*, 2027-2038.

Table 1. Blast analysis of *Ct*Lam81A with its homologues from family 81.

Organism	PDB ID	Query coverage (%)	Identity (%)	e-Value	Total Score
*Bacillus halodurans*C-125	5T4A	100	42	0	553
*Thermobifidafusca*YX-ER1	-	96	43	0	570
Glycine max	-	83	24	4e-35	130
*Aspergillus fumigatus*Af293	-	71	26	5e-35	130
Saccharomyces cerevisiae YPA84	-	64	27	2e-30	116
Rhizomucormiehei	4K3A	44	27	2e-25	112

Table 2. Binding analysis of *Ct*Lam81A with Laminari-oligosaccharides.

Ligand	Binding free energy, ΔG (kcal/mol)	Residue making polar interactions	Residue making Hydrophobic interactions
Laminaribiose	-7.29	Glu515, Asp438	Tyr430, Gly431, Glu435, Asn437, Phe 441, Trp588
Laminaritriose	-8.24	Asn513,Asn437	Phe508, Glu435,Gly431, Ser503, Tyr430, Asp438, Trp558, Trp589, Asp438, Trp588, His442, Phe441, Glu515, Glu519, His442, Ile599
Laminaritetraose	-7.01	Tyr430, Trp588, Trp637	Trp589, Tyr353, Asp665, Trp364
Laminaripentaose	-11.47	Tyr363,Lys367, Tyr430,Asp438, Asp513,Glu519	Trp637,Asp665,Ile599,Leu602,Trp364,Glu515,Trp588,Gly431,Ala502,Phe441
Laminarihexaose	-10.18	Tyr363,Glu519, Asp438,Asn513, Tyr430	Trp588,Phe508,Asn437,Ala502,Glu515,Phe441,Leu602,Lys367,Trp589,Asn590,Asn636,Trp637

Figures Legends

Figure 1: Secondary structure determination of *Ct*Lam81A by PsiPred server showing the amino acid residues involved in forming α-helix (cylinders), β-sheet (arrow) and random coil (continuous line).

Figure 2: Circular dichroism (CD) spectrum of *Ct*Lam81A for determining the percentage of secondary structure elements. The CD data were presented by difference in molar extinction coefficients (Δε, decilitre mol^{-1} cm^{-1}) as a function of wavelength. The purified *Ct*Lam81 at a concentration 0.7 μM in 50 mM phosphate buffer, pH 7.5 was used for CD analysis. The CD spectrum was recorded at 25°C using 1 nm bandwidth over far UV region between 190 to 250 nm at scanning rate of 50 nm/min.

Figure 3: (A) Cartoon representation of modeled structure of *Ct*Lam81A displaying N-terminal β-sandwich domain I (red colour), an $(\alpha/\alpha)_6$barrel domain II (green colour) and a short β-sandwich domain III (yellow colour) at C-terminal. Quality assessment of modelled *Ct*Lam81A by(B) Ramachandran plot of modelled *Ct*Lam81A, (C) Verify-3D and (D) Prosa web server.

Figure 4: (A) **S**uperimposition of *Ct*Lam81Amodeled structure (green) with *Bh*GH81 crystal structure (cyan) and (B)Superimposition of active site residues of *Ct*Lam81A (red colour) with active site residues of *Bh*GH81 (yellow colour) by molecular visualization system PyMol 2.

Figure 5: Multiple sequence alignment of *Ct*Lam81A with the homologuos protein of the family 81 glycoside hydrolase *viz. Bacillus halodurans* GH81 (*Bh*GH81; Q9KG76), *R. miehei* Lam81 (*Rm*Lam81A; KC847083), *T. fusca* Lam81A (*Tf*Lam81A, AAZ56163), *A. fumigatus* EngA (*Af*EngA; AAF13033), *S. cerevisiae* Eng2 (*Sc*Eng2; AAB82378) and *G. max*Gbp (*Gm*Gbp; BAA11407). The

conserved amino acid residues are shown in red background and semi conserved residues are shown in box. This figure was generated by EsPript3.0 (http://espript.ibcp.fr/).

Figure 6: Determination of hydrolytic mechanism (Retaining or Inverting) of *Ct*Lam81A enzyme by measuring the distance (from C_δ to C_δ) between the acid catalyst (Asp438) and catalytic base (Glu515) of *Ct*Lam81Aby molecular visualization system PyMol 2.

Figure7: Molecule dynamic *(*MD) simulation of modeled *Ct*Lam81A showing (A) RMSD plot and (B) Radius of gyration plot

Figure8: (A) Surface view of *Ct*Lam81A showing docking of Laminaripentaose with the amino acid residuesof the catalytic cleft and (B) 2D Schematic presentation ofLaminaripentaose interaction with the amino acid residues at active site of *Ct*Lam81A. Dashed lines show the hydrogen bonds and the amino acid residues shown in arc with spokes are making hydrophobic interactions with the ligand.

Fig. 1

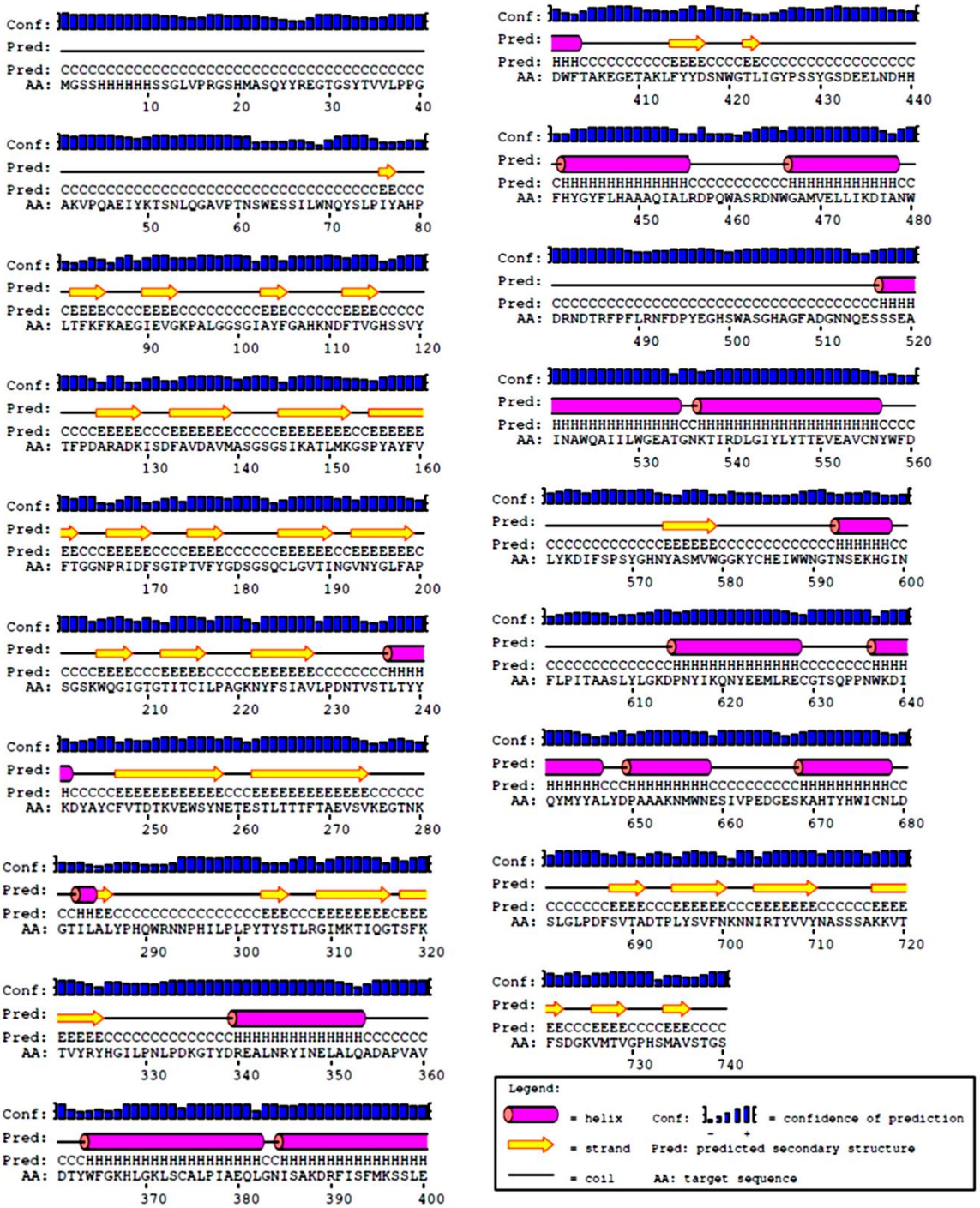

Conf:
Pred:
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
AA: MGSSHHHHHHSSGLVPRGSHMASQYYREGTGSYTVVLPPG
10 20 30 40
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCEECCC
AA: AKVPQAEIYKTSNLQGAVPTNSWESSILWNQYSLPIYAHP
50 60 70 80
Pred: CEEEECCCCEEEECCCCCCCCCEEECCCCCCEEEECCCCC
AA: LTFKFKAEGIEVGKPALGGSGIAYFGAHKNDFTVGHSSVY
90 100 110 120
Pred: CCCCEEEEECCCEEEEEEECCCCCEEEEEEEECCEEEEEE
AA: TFPDARADKISDFAVDAVMASGSGSIKATLMKGSPYAYFV
130 140 150 160
Pred: EECCCEEEEECCCCEEEECCCCCCEEEEEECCEEEEEEEC
AA: FTGGNPRIDFSGTPTVFYGDSGSQCLGVTINGVNYGLFAP
170 180 190 200
Pred: CCCCEEEECCCEEEEECCCCCEEEEEEECCCCCCCCHHHH
AA: SGSKWQGIGTGTITCILPAGKNYFSIAVLPDNTVSTLTYY
210 220 230 240
Pred: HCCCCCEEEEEEEEEEECCCEEEEEEEEEEEECCCCCC
AA: KDYAYCFVTDTKVEWSYNETESTLTTTFTAEVSVKEGTNK
250 260 270 280
Pred: CCHHEECCCCCCCCCCCCCCCCEEECCCEEEEEEECEEE
AA: GTILALYPHQWRNNPHILPLPYTYSTLRGIMKTIQGTSFK
290 300 310 320
Pred: EEEEECCCCCCCCCCCCCCHHHHHHHHHHHHCCCCCCC
AA: TVYRYHGILPNLPDKGTYDREALNRYINELALQADAPVAV
330 340 350 360
Pred: CCCHHHHHHHHHHHHHHHHHHCCHHHHHHHHHHHHHHHH
AA: DTYWFGKHLGKLSCALPIAEQLGNISAKDRFISFMKSSLE
370 380 390 400
Pred: HHHCCCCCCCCCCEEEECCCCEECCCCCCCCCCCCCCCC
AA: DWFTAKEGETAKLFYYDSNWGTLIGYPSSYGSDEELNDHH
410 420 430 440
Pred: CHHHHHHHHHHHHHHCCCCCCCCCCHHHHHHHHHHHCC
AA: FHYGYFLHAAAQIALRDPQWASRDNWGAMVELLIKDIANW
450 460 470 480
Pred: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCHHHH
AA: DRNDTRFPFLRNFDPYEGHSWASGHAGFADGNNQESSSEA
490 500 510 520
Pred: HHHHHHHHHHHHHHCCHHHHHHHHHHHHHHHHHHHCCCC
AA: INAWQAIILWGEATGNKTIRDLGIYLYTTEVEAVCNYWFD
530 540 550 560
Pred: CCCCCCCCCCCCCEEEEEECCCCCCCCCCCCCHHHHHCC
AA: LYKDIFSPSYGHNYASMVWGGKYCHEIWWNGTNSEKHGIN
570 580 590 600
Pred: CCCCCCCCCCCCCCHHHHHHHHHHHHHHCCCCCCCCHHHH
AA: FLPITAASLYLGKDPNYIKQNYEEMLRECGTSQPPNWKDI
610 620 630 640
Pred: HHHHHHCCCHHHHHHHHHCCCCCCCCCCHHHHHHHHHCC
AA: QYMYYALYDPAAAKNMWNESIVPEDGESKAHTYHWICNLD
650 660 670 680
Pred: CCCCCCCEEEECCCEEEEEECCCEEEEEEECCCCCCCEEEE
AA: SLGLPDFSVTADTPLYSVFNKNNIRTYVVYNASSSAKKVT
690 700 710 720
Pred: EECCCEEEECCCCEEECCCC
AA: FSDGKVMTVGPHSMAVSTGS
730 740
Legend:
= helix
= strand
= coil
Conf: = confidence of prediction
Pred: predicted secondary structure
AA: target sequence

Fig. 2

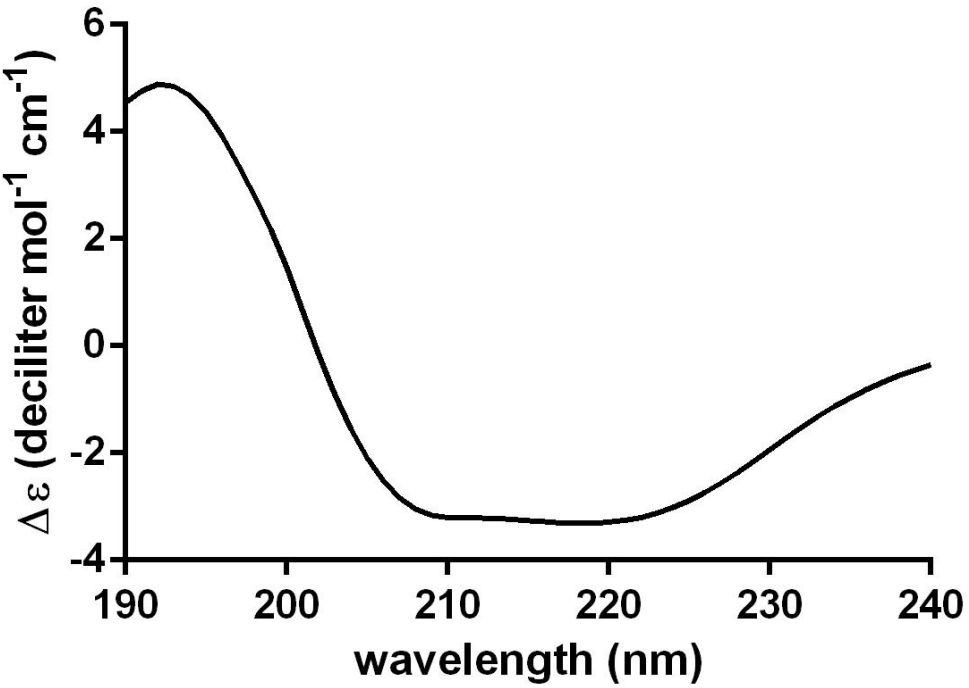
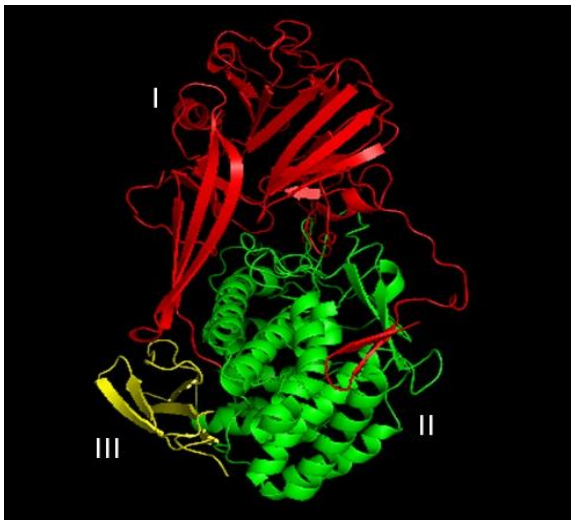
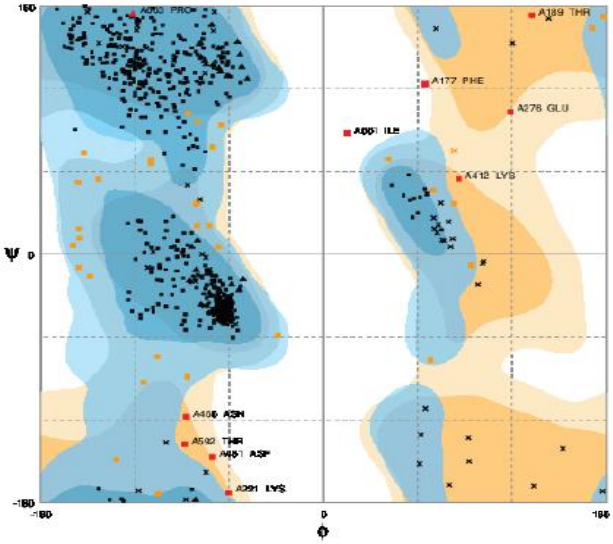

6
4
2
0
-2
-4
Δε (deciliter mol-1 cm-1)
190
200
210
220
230
240
wavelength (nm)

Fig. 3

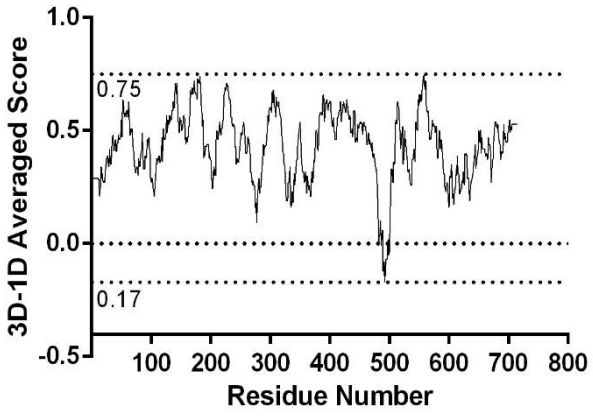
A


I
II
III

B


A189 THR
A177 PHE
A278 GLU
A412 LYS
Ψ
Φ

C


3D-1D Averaged Score
1.0
0.5
0.0
-0.5
0.75
0.17
100 200 300 400 500 600 700 800
Residue Number

D

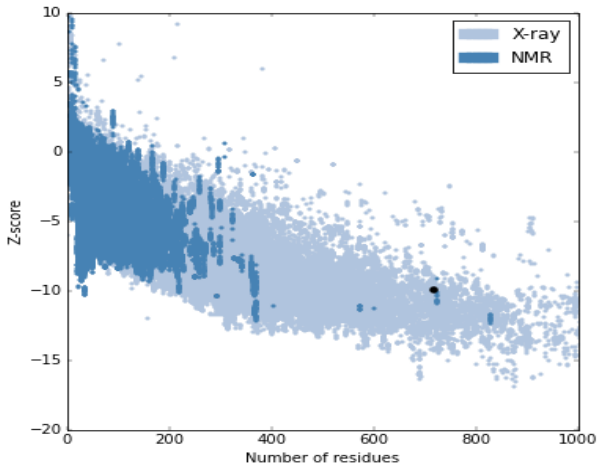

X-ray
NMR
Z-score
10
5
0
−5
−10
−15
−20
0 200 400 600 800 1000
Number of residues

Fig. 4

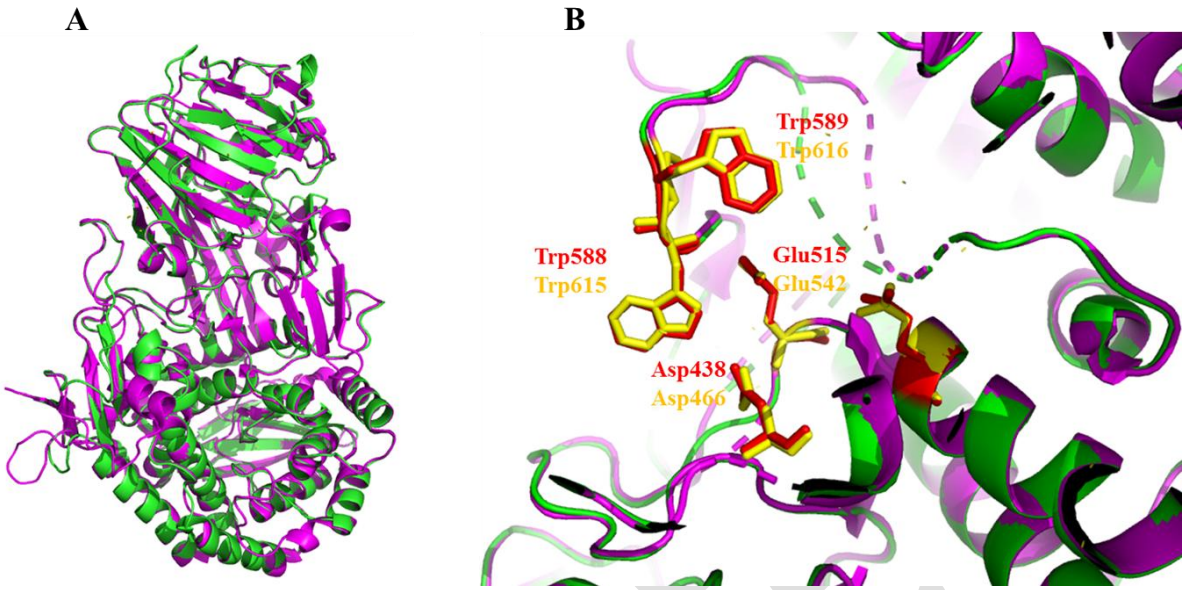
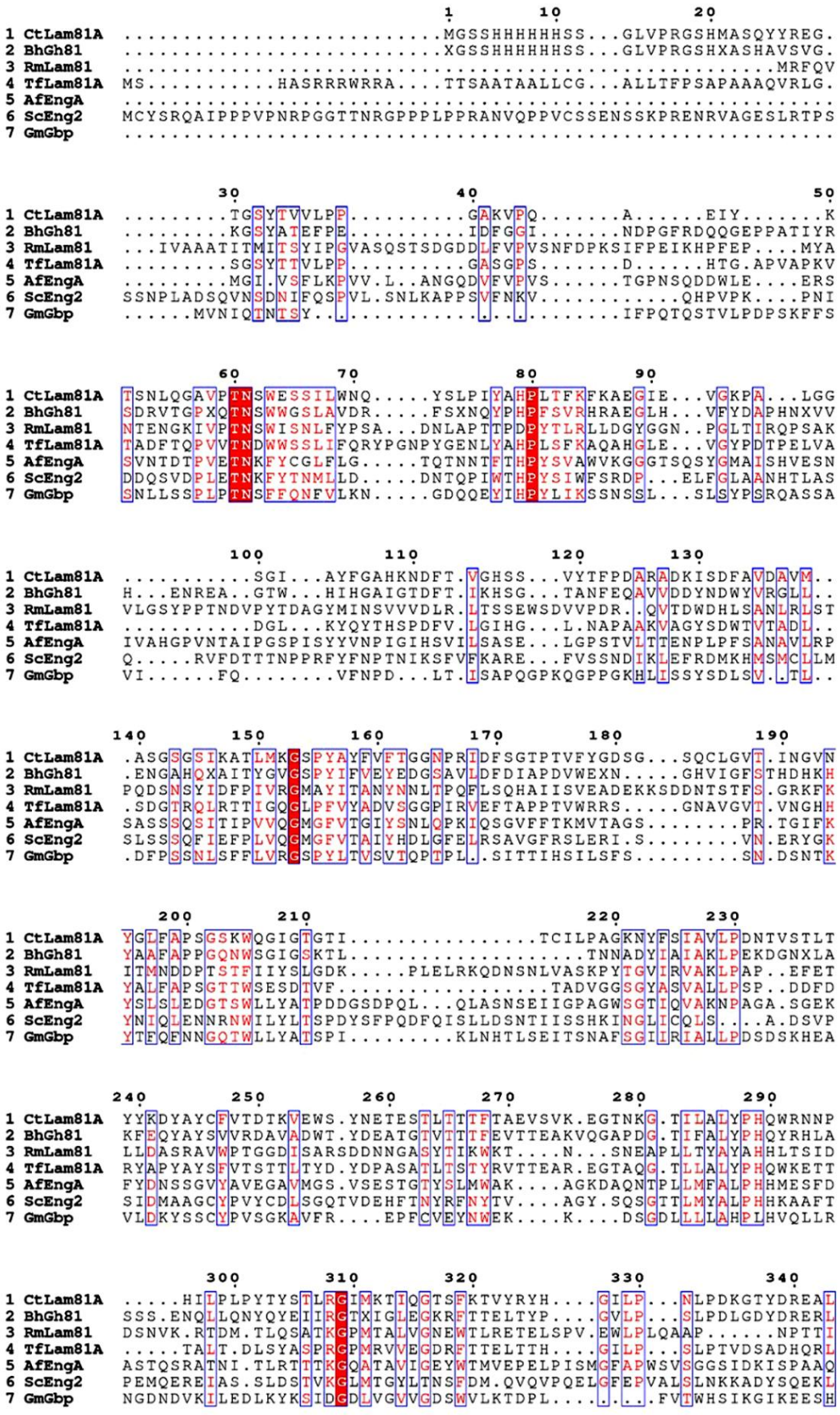
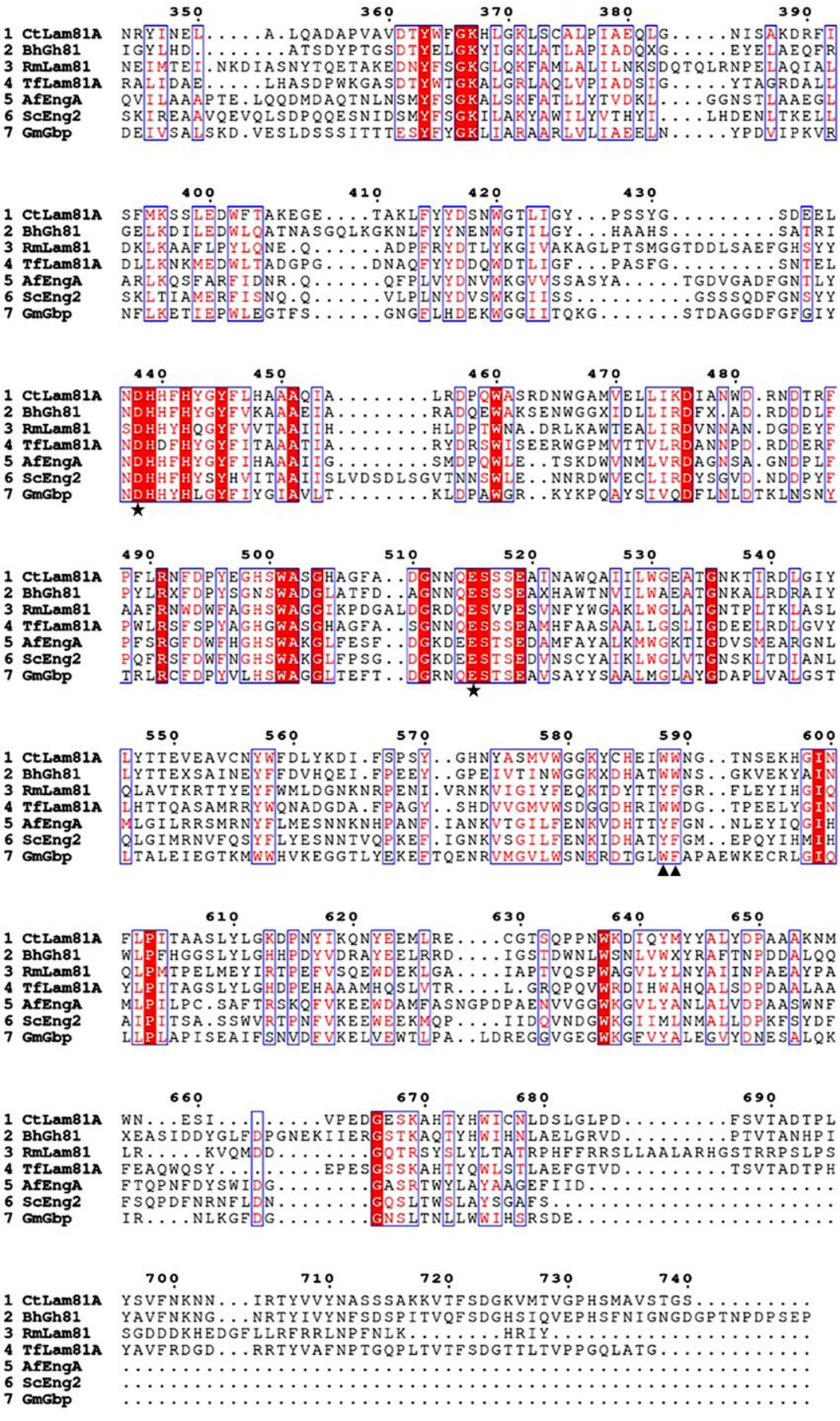


Fig. 5


1 CtLam81A
2 BhGh81
3 RmLam81
4 TfLam81A
5 AfEngA
6 ScEng2
7 GmGbp



Legend : ★ = Catalytic residues, ▲ = Binding residues

Fig. 6

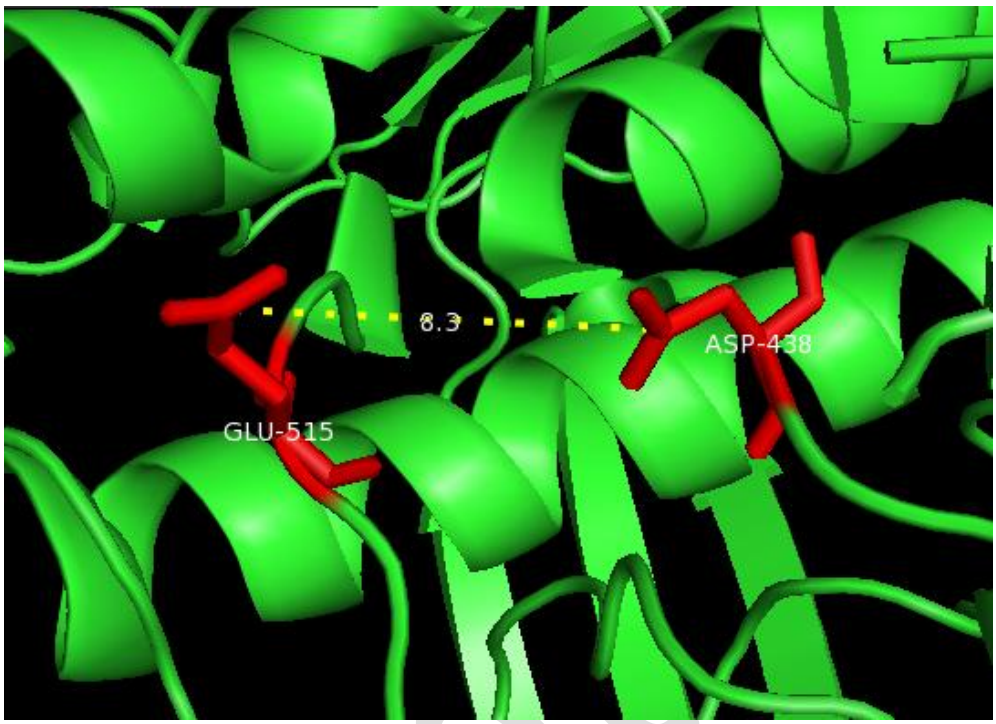


Fig. 7

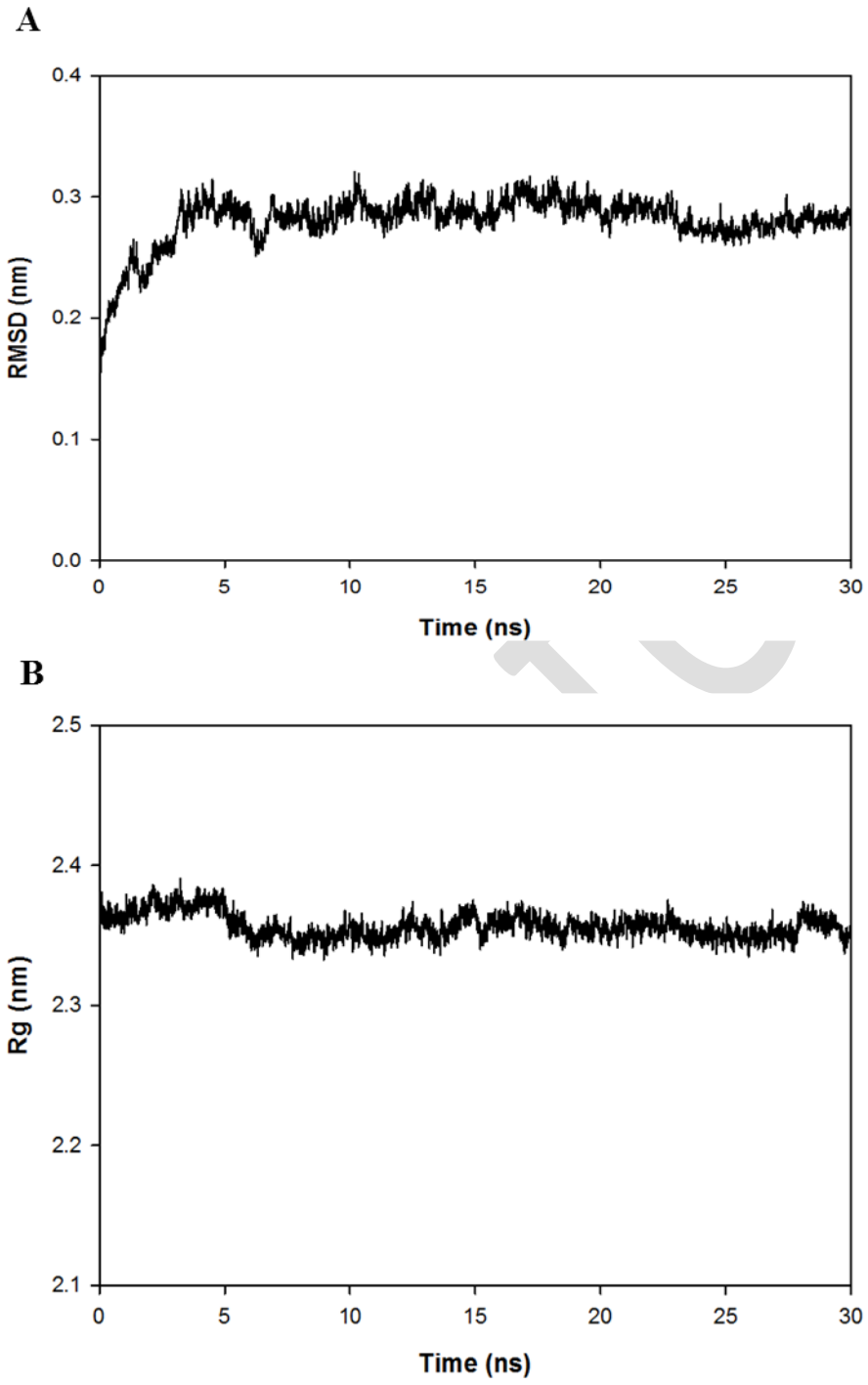
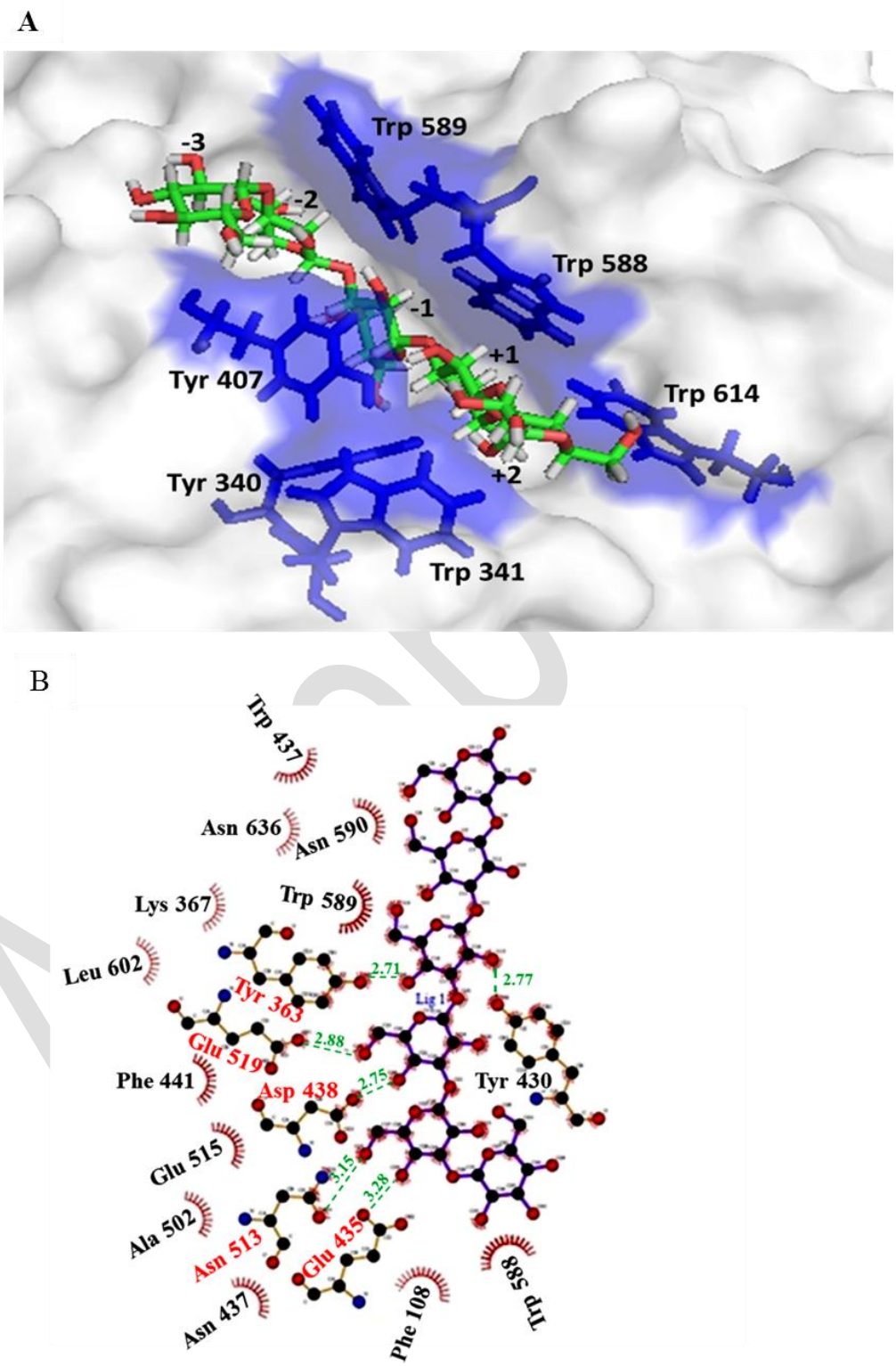

A
RMSD (nm)
0.0
0.1
0.2
0.3
0.4
0
5
10
15
20
25
30
Time (ns)
B
Rg (nm)
2.1
2.2
2.3
2.4
2.5
0
5
10
15
20
25
30
Time (ns)

Fig. 8


A
Trp 589
-3
-2
Trp 588
-1
+1
Tyr 407
Trp 614
+2
Tyr 340
Trp 341
B
Trp 437
Asn 636
Asn 590
Lys 367
Trp 589
Leu 602
2.71
2.77
Tyr 363
Lig 1
Glu 519
2.88
Phe 441
Asp 438
2.75
Tyr 430
Glu 515
3.15
3.28
Ala 502
Asn 513
Glu 435
Asn 437
Phe 108
Trp 588