Journal of Zhejiang University SCIENCE B
ISSN 1673-1581 (Print); ISSN 1862-1783 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Identification of novel catalytic features of endo-β-1,4-glucanase produced by mulberry longicorn beetle *Apriona germari*

SAMI Amtul Jamil†, HAIDER Mohammed Kamran

(*Institute of Biochemistry and Biotechnology, University of the Punjab, Lahore 54594, Pakistan*)

†E-mail: amtuljamilsami@yahoo.com; amtul@bb.pu.edu.pk

Received July 30, 2007; revision accepted Sept. 7, 2007

Abstract: Mulberry longicorn beetle, *Apriona germari*, has been reported to produce two endo-β-1,4-glucanases or AgEGases (accession Nos. Q6SS52 and Q5XQD1). AgEGase sequence contains catalytic motif (amino acid residues 37~48), which is the characteristic of family Glycohydrolase 45 and is identified as the substrate binding site. The application of bioinformatics approaches includes sequence analysis, structural modeling and inhibitor docking to relate the structure and function of AgEGases. We have dissected the sequence and structure of AgEGase catalytic motif and compared it with crystal structure of *Humicola insolens* endoglucanases V. The results show an involvement of sulfur containing amino acid residues in the active site of the enzyme. Cys residues and position of disulfide bonds are highly conserved between the two structures of endoglucanases of *A. germari*. Surface calculation of AgEGase structure in the absence of Cys residues reveals greater accessibility of the catalytic site to the substrate involving Asp42, a highly conserved residue. For the inhibition study, tannin-based structure was docked into the catalytic site of AgEGase using ArgusLab 4.0 and it resulted in a stable complex formation. It is suggested that the inhibition could occur through formation of a stable transition state analog-enzyme complex with the tannin-based inhibitor, as observed with other insect cellulases in our laboratory.

Key words: Cellulases, AgEGase, *Apriona germari*, Inhibition, Anthocyanidins
doi:10.1631/jzus.2007.B0765 **Document code:** A **CLC number:** Q55

INTRODUCTION

Cellulases are cellulose-degrading enzymes with a great potential to convert cellulosic material into its subunit-glucose (Beguin and Aubert, 1994). Cellulose hydrolysis by the enzymes is a major step in the carbon cycle on the globe (Coughlan and Ljungdhal, 1988; Betts *et al*., 1992). Organisms including microbes and lower animals are able to produce cellulases for the hydrolysis of cellulose (Bhat, 2000; Marsden and Gray, 1986; Sami *et al*., 1988; Watanabe and Tokuda, 2001). Endogenous animal cellulase genes from plant-parasitic nematodes and a termite have been identified and reported (Lee *et al*., 2004; 2005; Li *et al*., 2005; Watanabe and Tokuda, 2001; Watanabe and Sugimura, 2003). Since the discovery of insect cellulases there have been reports on the characteristics of cellulases. Comparisons between bacterial and insect cellulases on the basis of heterogeneity have been previously reported (Sami and Shakoori, 2006; Sami *et al*., 2005). However, there is presently little information available about the biochemical properties of pest cellulases.

Pests are the major cause of damage and loss of agricultural products, and the cellulase is the culprit behind it. To effectively control pests it is required to investigate the molecular basis of this damage. We have reported the structure-function relationship in *Apriona germari* using bioinformatics approaches (Sami *et al*., 2006). The cellulase of *A. germari* contains higher percentages of sulfur containing amino acid residues (Cys) (Lee *et al*., 2004; 2005). Reducing agents are generally considered for reducing the biological activity of hydrolytic enzymes. Computational approaches to analyze the important aspects of enzyme activity are very useful; we can find structural

bases for the various properties of an enzyme by identifying catalytic motif and the orientation of catalytic residues within a 3D structure. Homology modeling gives out theoretical models of those proteins for which suitable templates exist. These models can be further used for studies related to structure-function relationships. Endoglucanases produced by *A. germari* Q6SS52 and Q5XQD1 were used as models for the computational approach to analyze the important aspects of the enzyme activity, which could be further related to the structure-function relationship of cellulose enzyme.

MATERIALS AND METHODS

To generate a structural modeling of AgEGase, DeepView-SwissPDB Viewer (GlaxoSmithKline R & D), a computer program that provides a user friendly interface allowing to analyze several proteins at the same time, was used. The theoretical structure for AgEGase (endoglucanases of *A. germari*) was generated based on homology modeling using available sequences of *Trichoderma reesi*, *Fusarium oxysporum*, *Humicola insolens* and *Apriona germari*. DeepView submitted a target sequence to the ExPASy (**Ex**pert **P**rotein **A**nalysis **Sy**stem) proteomics server of the Swiss Institute of Bioinformatics (SIB), where BLAST was used to search the ExPDB (**Ex**pert **P**rotein **D**ata **B**ank) database for appropriate templates. *Humicola insolens* endoglucanase V was used as a template, as it shares 50% sequence similarity with AgEGase and belongs to the same family Glycohydrolase. After completing sequence alignment and its manual refinement, the project was submitted to Swiss-Model, an automated homology modeling server developed within the SIB, to produce a model. The model was then further refined by using Fit Selected Side Chains function of DeepView. The highly conserved residues including the family Glycohydrolase 45 (GH45) motif were highlighted in a 3D structure through SwissPDB Viewer. The orientation of side chains involved in catalysis was also compared, and the disulfide bonds related to the catalytic site were determined. SwissPDB Viewer computes molecular surface through an algorithm that defines molecular surface as an area that can be reached with the surface of a solvent molecule (with a radius of 1.4 Å) rolled over the protein. The catalytic motif of AgEGase along with active site residues was docked with the inhibitor using the ArgusLab 4.0, a molecular modeling, graphics and drug design program (Planaria Software LLC, Seattle, WA, USA) (Thompson, 2004). The best orientation of the enzyme molecule during inhibitor binding was selected with the score −30.89 kJ/mol within the binding site grid of dimensions of 10 Å×10 Å×11.239 Å.

RESULTS

A multiple sequence alignment of AgEGase amino acid sequence with other sequences from the same family including endoglucanases of *Trichoderma reesi*, *Humicola insolens* endoglucanase V, *Fusarium oxysporum* and *Apriona germari* indicated the position of catalytic motif and other important residues with observed catalytic roles (Fig.1). The homology was observed at the N-terminus which forms the active site of the enzyme in all four structures studied. On the basis of this information the model of cellulases of *A. germari* was built. The template used for the prediction of the structure was the experimentally determined structure of endoglucanase V of *H. insolens* (Fig.2), which was reported by Davies *et al*.(1993). The theoretical results show that the prominent part of the globular structure is the six-stranded α-barrel domain, while the most prominent part in the globular structure of AgEGase is the six-stranded β-barrel domain. A deep groove runs through the enzyme surface and virtually separates the β-barrel and loops, some of which contain short helices. The characteristic catalytic motif (amino acids 37~48 highlighted in green in Fig.2) resides in the groove on the protein surface. The groove actually forms the substrate binding site as it has been determined for endoglucanase V (Davies *et al*., 1993) and it is a series of at least six hexose ring binding sub-sites. This groove forms the catalytic motif (Fig.3). Another important aspect of AgEGase activity is the role of disulfide bonds. The highly conserved residues including GH45 (Henrissat, 1991; Henrissat and Bairoch, 1993) motif were highlighted in a 3D structure generated by SwissPDB Viewer. The orientations of side chains involved in catalysis were also compared. The disulfide bonds related to

catalytic site were determined to be Cys191-Cys202, Cys167-Cys43 and Cys64-Cys98. We found that the enzyme activity increased at low mercaptoethanol concentrations. However, it did not persist as the concentration was raised (data not included). We suggest that the increase in the enzyme activity was resulted from a reduction of disulfide bonds present at the catalytic motif including Cys191-Cys202, Cys167-Cys43 and Cys64-Cys98 (Fig.3). These results, when correlated with structural modeling, give an important and interesting insight about the active site of AgEGase. The reduction of those disulfide bonds which are associated with the active site allows a greater accessibility to the substrate. In Fig.4a, AgEGase molecular surface is shown in the presence of Cys residues and disulfide bonds. A dark blue cavity marks the catalytic cleft, which indicates that it is buried deep into the protein molecule. On the other hand, molecular surface was modeled in the absence of Cys residues (Fig.4b) and hence no disulfide bonds were formed. The surface calculated reveals cavities and exposed molecular surfaces whose surface area can be examined in order to define the binding sites and catalytic clefts. The deep catalytic cleft is exposed to a significant extent in the absence of disulfide bonds. Moreover, a closer look at the exposed catalytic cavity shows that Asp42 also has an increased accessibility within the catalytic site in order to initiate the nucleophilic attack. Most importantly, the substrate can be more exposed to the active site, the Asp residue, which makes a nucleophilic attack on the β-glycosidic linkage. Dissection of the amino acid residues of AgEGases showed that the surface calculation of AgEGase structure in the absence of Cys residues reveals a greater accessibility of the catalytic site to the substrate involving Asp42, a highly conserved residue.

Grape leaves extract has been reported as an inhibitor of microbial and plant cellulases (Bell *et al*., 1965); flavonoid-like compounds present in grapes are responsible for inhibition of the enzyme (Fig.5). A bioinformatics approach was applied and the structure of the compound anthocyanidin was studied for docking the enzyme of AgEGase of *A. germari*. Anthocyanidin that is present in grape leaves extract was reported to be the inhibitor for cellulose hydrolyzing enzymes. For docking purpose we used part of anthocyanidin structure that is actively involved in inhibition. The catalytic motif of AgEGase along with the active site residues was docked with the inhibitor in ArgusLab 4.0 (an empirical scoring function for structure-based binding affinity prediction), selected the best candidate pose with the score −30.89 kJ/mol within the binding site grid of dimensions of 10 Å×10 Å×11.239 Å (Figs.5 and 6). For the inhibition study the tannin-based structure was docked into the catalytic site of AgEGase using ArgusLab 4.0. It resulted in a stable complex formation. It is suggested that the inhibition could occur through the formation of a stable transition state analog-enzyme complex with the tannin-based inhibitor as observed with other insect cellulases in our laboratory.

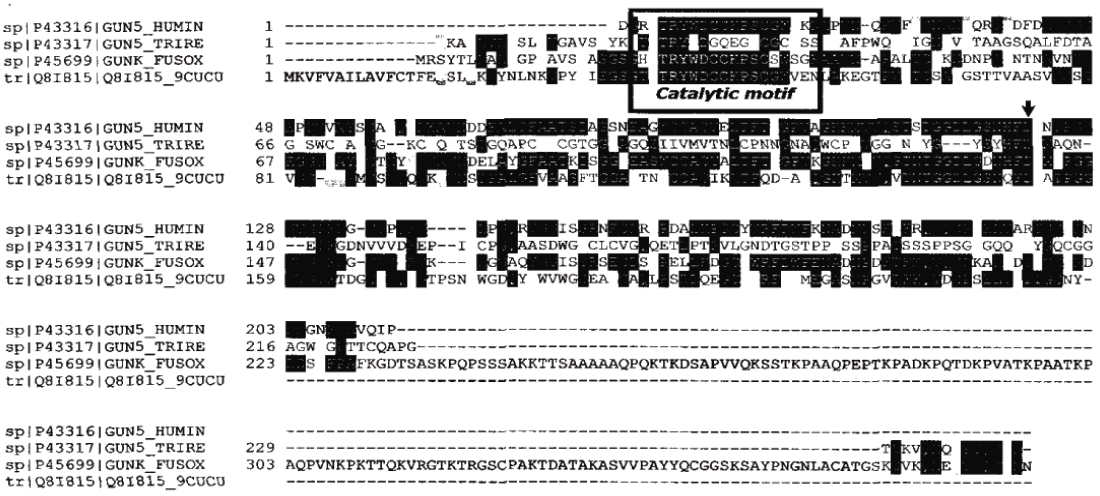
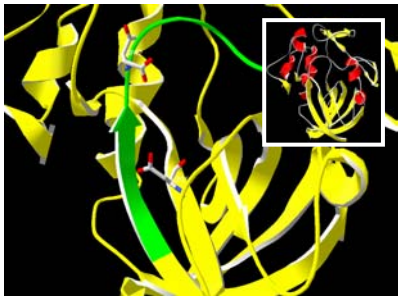
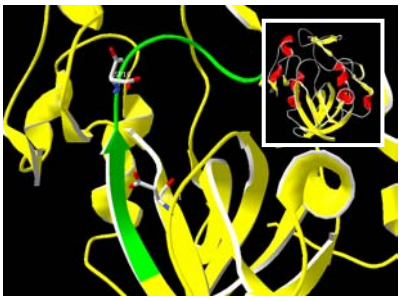


Fig.1 Multiple sequence alignment of AgEGase and other members from family Glycohydrolase 45
Block represents highly conserved catalytic motif. Arrow indicates another conserved residue with observed catalytic role. GUN5_HUMIN: *Humicola insolens*; GUN5_TRIRE: *Trichoderma reesi*; GUNK_FUSOX: *Fusarium oxysporum*; Q81815_9CUCU: *Apriona germari*

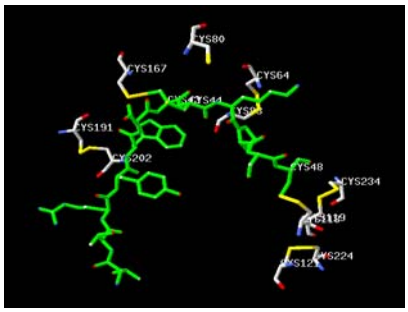


(a)

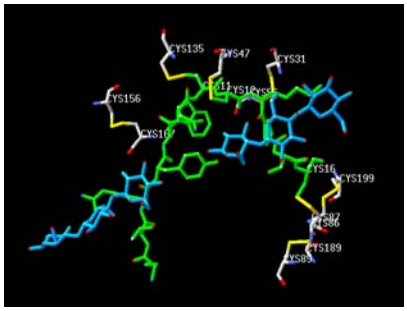


(b)

Fig.2 ***Apriona germari*** **cellulase I (a) and** ***Humicola insolens*** **endoglucanase V (b) (shown in the inset) are members of the same family Glycohydrolase 45. Catalytic site of both enzymes resides in the same motif and contains highly conserved residues which have active role in catalysis (ScanProsite pattern PS01140). This motif (highlighted in green) can be computationally assessed for inhibitor docking**

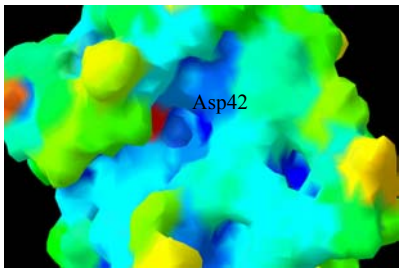


(a)

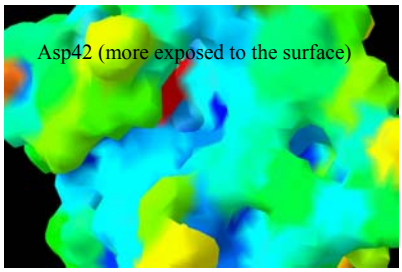


(b)

Fig.3 Disulfide bonds associated with active site in AgEGase (a) and endoglucanase V (b). The similar spatial orientation of disulfide bonds suggests that Cys have putative role in specifying the fold geometry that is characteristic of family Glycohydrolase 45 (part of cellulose, a substrate molecule, is shown in blue, catalytic motif in green and Cys in yellow)



(a)



(b)

Fig.4 Molecular surface of AgEGase (a) in the presence and (b) in the absence of disulfide bonds. Molecular surface is colored according to the accessibility; deep blue areas show buried cavities whereas light blue color indicates exposed molecular surfaces. Molecular surface with distinct properties is colored differently. After partial reduction there is a greater accessibility of catalytic site involving Asp42 (shown in red)

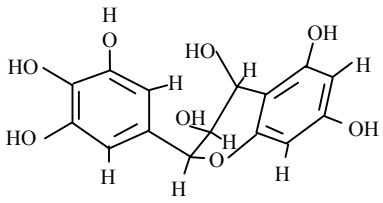


Fig.5 An anthocyanidin (condensed tannin molecule)

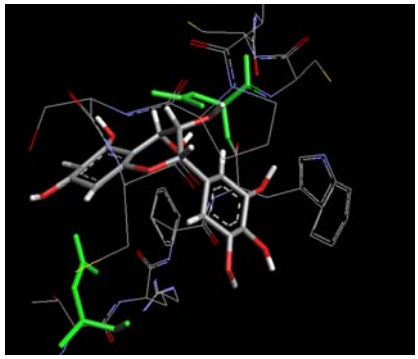


Fig.6 Docking the inhibitor (an anthocyanidin) to the binding site of AgEGase in ArgusLab results in 11 unique configurations. The one with the highest score (scoring function Ascore) is shown here and fits well into the purposed model of enzyme inhibition. Catalytic residues are shown in green

DISCUSSION

Sequence annotations based on the biochemical information about enzymatic catalysis of a representative member of a protein family are quite useful in building models that can refine our knowledge about the mode of action and inhibition of other enzymes. Sequence motif characteristic of GH45 is useful in correlating several structure-based properties of the AgEGase such as an essentially conserved fold, an active site geometry and also the catalytic mechanism. A comparison of endoglucanases of *T. reesi*, *H. insolens* endoglucanase V, *F. oxysporum* and *A. germari* was performed (Fig.1). In order to predict the catalytic mechanism of AgEGase it is important to observe a spatial organization of conserved catalytic residues. The mechanism of action of endoglucanase V was first reported by Davies *et al*.(1993). They found that the catalysis involves Asp10 and Asp121 that are located in the deep groove. The reaction takes place by participation of a water molecule beside the other catalytic residues. The water molecule bound to Asp10 makes a nucleophilic attack on β-1,4-glycosidic linkage so that the cleavage occurs at the reducing end of the substrate and ensues inversion at C1. In the case of AgEGase, Asp42 and Asp152 occupy almost the same positions as do Asp10 and Asp121 in endoglucanase V, respectively (Fig.2). Hence, the catalytic fold geometry dictates the catalytic mechanism and it appears that the AgEGase catalyzes the breakdown of β-1,4-glycosidic linkage in the same manner. The optimum pH of the enzyme is in acidic range, which keeps the catalytic residue Asp42 in an ionized form so that it can bind water molecule for the nucleophilic attack. Asp152 acts to stabilize the oxonium ion that forms as a result of protonation of linking O-atom between D and E sugar residues. In fact, the AgEGase and endoglucanase V utilize multiple factors for a high efficient catalysis, including acid catalysis (nucleophillic attack by Asp42 bound water) and electrostatic and transition state catalysis (stabilization of D-ring oxonium ion). It was recorded that alignment patterns of conserved residues, especially those that are characteristic of the GH45 family, are similar. Catalytic motif in this family has been identified as a signature pattern (ScanProsite pattern PS01140), which is also shared by the AgEGase (Figs.1 and 2). Moreover, the conservation of aromatic residues and Cys residues at several positions also appears to be a strong characteristic feature of GH45. Since the positions of disulfide bonds are conserved in 3D structures, so the disulfide bonds are indicated to be involved in specifying the fold geometry especially the substrate binding cleft, besides maintaining an overall structural integrity of the enzyme. In further exploring the structural basis of these results, however, at the sequence level it appears that highly conserved Cys residues and aromatic residues that span a part of the catalytic cavity have a concerted role in maintaining a balance between the catalytic efficiency and permissible fold geometry that allows a close hydrophobic contact in the substrate binding cleft as well as the participation of polar catalytic residues in acid catalysis and stabilization of transition state. The role of highly conserved Cys residues and hence spatially conserved disulfide bonds is evaluated through sequence analysis and experiments in other insects (unpublished data). The role of disulfide bonds can be explained in terms of fold geometry and catalytic efficiency. Disulfide bonds specify the cleft geometry in such a way that upon substrate binding the active site utilizes all these mechanisms to achieve catalysis. Thus, the disulfide bonds that are associated with the active site strengthen the catalytic residues acting via multiple mechanisms. In the absence of disulfide bonds, although the catalytic activity increases, the exposed catalytic fold generated appears to compromise some favorable enzyme-substrate interactions that allow a multiple-catalysis strategy. Enzyme surface calculations in both the presence and absence of disulfide bonds reveal an increase in relative accessibility of catalytic cavity and catalytic residues in the later case. It is suggested that the disulfide bonds act as 'molecular clips' to determine catalytic fold geometry that allows efficient catalysis through simultaneous multiple mechanisms.

It was reported that a sugar residue bound at D-site undergoes a distortion that renders it in a planar half chair conformation (Schindler *et al*., 1977; Strynadka and James, 1991). In fact, this distortion is a requirement of the D-ring oxonium ion formed during catalysis, in order to achieve resonance stabilization, its atoms C1, C2, C5 and O5 must be coplanar. Consequently, the binding conformation of

D-ring oxonium ion is the reaction's transition state. The best candidate pose chosen by scoring function of docking program is the one with conformation and binding orientation that is quite similar to that of the D-ring oxonium ion. Carboxylate group of Asp42 most probably stabilizes the pyran ring of inhibitor through electrostatic interaction. Hydrozyl groups can form multiple hydrogen bonds at the binding site, utilizing full potential of binding cleft residues. The multiple weak interactions developed lead to the complex formation, which, in theory, binds the active site more stably because the inhibitor structure is analogous to the transition state of the reaction. Given the presence of similar catalytic motif in endoglucanase V and the AgEGase, the catalytic mechanism can be reliably predicted on the basis of a computational assessment of a natural inhibitor, and a model of enzyme inhibition is purposed for cellulases of the GH45, with a transition state analog-based inhibition It reveals the potential of natural substrate analog inhibitors (flavonoid) in blocking insect cellulase activity. Our model of AgEGase inhibition is based on the formation of the stable transition state analog-enzyme complex and is supported by the output of docking procedure. The crystallographic evidence of this model is required in order to evaluate the conformity of this model to the actual mechanism of AgEGase inhibition.

References

Beguin, P., Aubert, J.P., 1994. The biological degradation of cellulose. *Microbiol. Rev.*, **13**(1):25-58.

Bell, T.A., Eichells, J.I., Smarr, W.G.Jr, 1965. Pectinase and cellulase enzyme inhibition from sericea and certain other plants. *Botanical Gazette*, **126**(1):40-45. [doi:10.1086/336293]

Betts, W.B., Dart, R.K., Ball, A.S., Pedler, S.R., 1992. Biosynthesis and Structure of Lignocellulose. *In*: Betts, W.B. (Ed.), Biodegradation of Natural and Synthetic Materials. Springer-Verlag, Berlin, Germany.

Bhat, K.M., 2000. Cellulases and related enzymes in biotechnology. *Biotechnol. Adv.*, **18**(5):355-383. [doi:10.1016/S0734-9750(00)00041-0]

Coughlan, M.P., Ljungdhal, L.C., 1988. Comparative Biochemistry of Fungal and Bacterial Cellulolytic Enzyme Systems. *In*: Aubert, J.P., Beguin, P., Millet, J. (Eds.), Biochemistry and Genetics of Cellulose Degradation. Academic Press, London. p.11-30.

Davies, G.J., Dodson, G.G., Hubbard, R.E., Tolley, S.P., Dauter, Z., Wilson, K.S., Hjort, C., Mikkelsen, J.M., Rasmussen, G., Schuelein, M., 1993. Structure and function of endoglucanase V. *Nature*, **365**(6444):362-364. [doi:10.1038/365362a0]

Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.*, **280**:309-316.

Henrissat, B., Bairoch, A., 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.*, **293**:781-788.

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2004. cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **139**(1):107-116. [doi:10.1016/j.cbpc.2004.06.015]

Lee, S.J., Kim, S.R., Yoon, H.J., Kim, I., Lee, K.S., Je, Y.H., Lee, S.M., Seo, S.J., Dae Sohn, H., Jin, B.R., 2005. A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: gene structure, expression, and enzymatic activity. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.*, **140**(4):551-560. [doi:10.1016/j.cbpc.2004.12.003]

Li, Y., Guo, R., Yini, Q., Dingi, M., Zhang, S., Xui, G., Zahoi, F., 2005. Purification and characterization of two endo glucanases from Mollusca *Ampullaria crossean*. *Acta Biochimica et Biophysica Sinica*, **37**(10):702-708. [doi:10.1111/j.1745-7270.2005.00099.x]

Marsden, W.L., Gray, P.P., 1986. Enzymatic hydrolysis of cellulose in lignocellulosic materials. *CRC Crit. Rev. Biotech.*, **3**:235-276.

Sami, A.J., Shakoori, A.R., 2006. Heterogeneity in cellulases of some of the local agricultural insects pest. *Pak. J. Zool.*, **38**(4):337-340.

Sami, A.J., Akhter, M.W., Malik, N.N., Naz, B.A., 1988. Production of free and substrate bound cellulases of *Cellulomonas flavigena*. *Enzyme Microbial Technol.*, **10**(10):626-631. [doi:10.1016/0141-0229(88)90110-X]

Sami, A.J., Rehman, F.U., Alam, M., Haider, R., 2005. A Repoprt on Multiplicity of Cellulases in Pests and Its Comparison with Multiple Forms of Bacterial Cellulases. *In*: Biocatalysis: Enzymes Mechanism and Bioprocess. Biochemical Society Focussed Meeting Manchester, 21−22 Nov, Poster No. 14.

Sami, A.J., Haider, K.M., Rehman, F.U., 2006. Determination of Structure-Function Relationship in *Apriona germari* Endoglucanases Using Bioinformatics Approaches. Biochem. Soc. Meeting Biosci., 23-27 July Glasgow Abstract No. 558.

Schindler, M., Assaf, Y., Sharon, N., Chipmann, D.M., 1977. Mechanism of lysozyme catalysis: role of ground-state strain in sub-site D in hen egg white and human lysozyme. *Biochemistry*, **16**(3):423-431. [doi:10.1021/bi00622a013]

Strynadka, N.C.J., James, M.N.G., 1991. Lysozyme revisted: crystallographic evidence for distortion on an N-acetyl muramic acid residue bound in site D. *J. Mol. Biol.*, **220**(2):401. [doi:10.1016/0022-2836(91)90021-W]

Thompson, M., 2004. ArgusLab 4.0 Poster Presentation: Molecular Docking Using ArgusLab: An Efficient Shape-Based Search Algorithm and the Ascore Scoring Function. Fall 2004 ACS Meeting, Philadelphia.

Watanabe, H., Tokuda, G., 2001. Animal cellulases. *Cell Mol. Life Sci.*, **58**(9):1167-1178. [doi:10.1007/PL00000931]f

Watanabe, H., Sugimura, M., 2003. Evidence of the presence of bilaterian animals. *Proc. Biol. Sci.*, **270**(3):69-70.