

Original article

## Computational structure prediction and analyze active ligand binding site of defense and lytic enzymes of *Trichoderma harzianum*

Ved Ratan, Supriya Dixit, Mukesh Srivastava\*, Shubha Trivedi, Abhishek Mishra, Y.K. Srivastava and D.K. Srivastava\*\*  
 Biocontrol Lab, Department of Plant Pathology, Chandra Shekhar Azad University of Agriculture and Technology,  
 Kanpur-208002, Uttar Pradesh, India

\*Registrar, Rani Laxmi Bai Central Agricultural University, Jhansi-284003, Uttar Pradesh, India

\*\*Joint Director, Council of Science and Technology, Bans Mandi, Qaiserbagh, Lucknow-226018, Uttar Pradesh, India

Received June 13, 2018; Revised July 30, 2018; Accepted August 5, 2018; Published online December 30, 2018

### Abstract

*Trichoderma* spp. are considered as efficient biocontrol agent that can significantly reduces the growth of several soil borne and plant pathogens such as *Rhizoctonia solani*, *Sclerotium rolfsii*, *Phythium aphanidermatum*, *Fusarium oxysporum* in several mode of action. *Trichoderma harzianum* is one the important species in genus *Trichoderma*, which is capable of producing several effective lytic enzymes and antifungal antibiotics that compete to other fungal pathogens and promotes plant growth. The aim of present study is to predict and analyze the tertiary structure and their potential binding sites through bioinformatic tools and techniques. The protein sequences of enzymes were retrieved from UniProt database, followed by modelling of tertiary structure by Swiss-model Workspace and further validated using PROCHECK server which showed that from 75% to 90% residues are in favored region of Ramchandran Plot. These different validation steps proved that the predicted models are stable and it will provide an insight to its functional aspect which is based on tertiary structures. Furthermore, functional and conserved motifs were predicted through PROSITE database. These findings allow us to determine the protein families and domain which remain conserved throughout the evolution which may act as inducing or suppressing the biological activity of protein. Ligand binding site of enzymes has been predicted using SiteHound server by using four different chemical probes which allow us studying different ligand binding site. Thus, this study supported a scientific base for 3D structure modelling of lytic and defence enzymes and opens the new opportunities for further investigations in biological control of phytopathogens.

**Key words:** Physicochemical properties, signal peptides, transmembrane region, comparative modelling, ligand binding site

### 1. Introduction

The modern agriculture approaches increases the use of genetically modified crop plants, fertilizer, pesticides, various biotechnologically advanced equipment and water to produce large quantity of single modified crop which causes increase in disease susceptibility which will lead to destructive crop diseases caused by microorganism (Waard *et al.*, 1993). Therefore, nowadays, various researches have been aimed to control plant diseases by replacing the pesticides with eco-friendly biological control agents which could be rhizospheric fungi and bacteria, having antagonistic activities towards pathogens. Biological control agents should have ability to survive under unfavorable conditions, efficient to utilize nutrient and increases the quality of soil (Haran *et al.*, 1996).

From decades, *Trichoderma* species are well known saprophytic fungi habitat in the rhizosphere, which have received significant

observation as potential fungal antagonist that could be used as effective biopesticides. Among all known *Trichoderma* spp (100), *Trichoderma harzianum* was found to be very efficacious bioagent against soil borne phytopathogens (Green *et al.*, 1999).

In plants combination of dynamic mycoparasitism and effective induced defense strategies revealed significant effect of *T. harzianum*. During mycoparasitism attack, *Trichoderma* releases various hydrolytic enzymes such as chitinases, glucanases, cellulase, xylanase and protease which degrade pathogenic fungal cell wall and utilize its nutrient (Strakowska *et al.*, 2014). *T. harzianum* also produces another group of defence and antioxidant enzymes which includes catalase, peroxidase (PO), polyphenoloxidase (PPO), PR proteins, phytoalexins, chalcone synthase, phenols and phenylalanine-ammonia lyase (PAL); these enzymes are accumulated in root which enhance the defence mechanism in plant resistance process (Fotoohiyan *et al.*, 2015). Peroxidases played important role in plant disease resistant by oxidizing a variety of phenolic compounds and implicated those compounds in many different physiological processes such as incorporation of phenolics into the cell wall, lignification, wound healing, and pathogen defense (Khaled *et al.*, 2008). Chakraborty and Chatterjee (2007) studied the activity of peroxidase (PO) and polyphenol oxidase (PPO) of

**Author for correspondence:** Mrs. Supriya Dixit  
 Biocontrol Lab, Department of Plant Pathology, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur-208002, Uttar Pradesh, India

**E-mail:** supriya.dixit.28@gmail.com

**Tel.:** +91-7522090030

Copyright © 2018 Ukaaz Publications. All rights reserved.

Email: ukaaz@yahoo.com; Website: www.ukaazpublications.com

brinjal plants during *T. harzianum* and *F. solani* interaction to find any correlation that may exist in host defense mechanism and they noticed that *T. harzianum* was significantly interacts with peroxidase activity which is increased in the period of infection and lasts till the end of incubation period which also indicates that there is a correlation with resistance of plant against infection. Thus, the aim of this study is to explore more about the lytic and defence enzymes through various computational software and techniques. This goal requires the knowledge of three dimensional structures of proteins obtained through X-ray crystallography and NMR spectroscopy. The crystal structures of defense and lytic enzymes are not available in protein data bank (PDB), but these experimental techniques are very expensive, time consuming and complex process. Hence, by using computational technique, the 3D structure of defense and lytic proteins has been modeled through comparative or homology modelling which predict model based on experimental 3D structure of related homologous protein (serve as template). The Swiss model server is used to construct the three dimensional structure of proteins and, further validated through Web server software by plotting Ramachandran Plot.

However, the studies suggested that approximate by one third of proteins are associated with small probes and their presence and absence directly affects the catalytic mechanism of proteins and it is also helpful to stabilize the tertiary structure of proteins (Shi and Chance, 2008). So, in order to identify the most potential probe with their putative ligand binding site, the SITEHOUND-Web server (<http://sitehound.sanchezlab.org>) is used.

## 2. Materials and Methods

The protein sequences of defense and lytic enzymes produced by *T. harzianum* were retrieved from UniProt Database in fasta format and used further for the prediction of physicochemical properties like theoretical pI, molecular weight, half-life, instability index and grand average of hydropathicity through ExPASy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (Wilkins *et al.*, 1999). The presence of signal peptides in defence and lytic proteins was analysed by using SignalP 4.1 server (Petersen *et al.*, 2011). The protein sequences were also used for the prediction of conserved motif which plays an important

role for inducing defense in plants and also performs catalytic activity. Thus, ScanProsite program (De Castro *et al.*, 2006) was used which allows scanning of protein sequence for matches against the PROSITE collection of motifs with a meaningful signature or pattern and then similarity searching was done for validation of the identified motif in other *Trichoderma* species in order to check whether the identified motif is present in other strains of *Trichoderma* or not. For comparative modelling, Swiss model Automated Workspace (Arnold *et al.*, 2006) was used to determine the 3D structure. The predicted models were further evaluated in order to know the quality of structures. PROCHECK server was used to validate predicted structure. The PROCHECK server generates 'Ramachandran Plot' on the basis of stereochemical configuration of amino acids (Laskowski *et al.*, 1993). Identification of binding site with favorable interaction with four different ligands (Methyl Carbon, Phosphate Oxygen, Hydroxyl and Aromatic) were done using SITEHOUND Web server (Hernandez *et al.*, 2009).

## 3. Results and Discussion

The present computational studies focused on sequence and structural analysis of lytic and defence enzymes which participates in fungal cell wall degradation includes chitinases, glucanases, protease and peroxidase of *T. harzianum*. Sequences of lytic enzymes were retrieved from UniProt Database and are varied in length from 300 to 1040 amino acid (Table 1) and used for further analysis.

**Table 1:** The pathogenic proteins name and their accession numbers

Enzymes	Accession number (Protein sequence)
Alkaline Proteinase	A0A0F9X8B4
Beta-1,3 exoglucanase	O14402
Beta-1,3-glucan-binding	A0A0F9Y010
Beta-1,6-glucanase	B9VQ17
Catalase-peroxidase	A0A0F9X3Z8
Endochitinase 33	Q12713
Endochitinase 37	Q8NJQ5
Endochitinase 42	P48827

**Table 2:** Prediction of theoretical physicochemical properties for defence and lytic enzymes

	Alkaline proteinase	Beta-1,3 exo - glucanase	Beta-1,3-glucan -binding	Beta-1,6-glucanase	Catalase-peroxidase	Endo-chitinase 37	Endo-chitinase 37	Endo-chitinase 42
<b>No. of AA</b>	408	1032	481	490	823	337	321	423
<b>MW</b>	43732.40	107912.51	53342.39	51862.30	88742.47	35478.19	34348.34	46056.52
<b>pI</b>	5.71	4.66	5.32	4.52	5.49	4.55	5.07	7.01
<b>Ext. coefficient</b>	50880	139715	125610	102010	146915	81610	66265	89730
<b>Instability index</b>	30.40	35.20	35.08	27.57	29.66	25.08	29.68	23.57
<b>Stability</b>	stable	stable	stable	stable	stable	stable	stable	stable
<b>Aliphatic index</b>	89.24	72.11	65.05	66.24	74.31	85.16	67.88	75.53
<b>GRAVY</b>	-0.055	-0.109	-0.553	-0.188	-0.254	0.104	-0.135	-0.311

### 3.1 Physicochemical analysis

By using the ProtParam tools, several physicochemical properties were computed for alkaline proteinase, Beta-1,3 exoglucanase, Beta-1, 3-glucan-binding protein, Beta-1,6-glucanase, Catalase-peroxidase, Endochitinase 33, Endochitinase 37 and Endochitinase 42 (Table 2).

The predicted instability index for *T. harzianum* enzymes are below 40 which indicated that all enzymes are in stable form. A protein with an instability index lower than 40 is predicted as stable, while a value greater than 40 predicts that the protein may be unstable. The significance for computation of instability index is to store proteins in correct solvent. Grasso *et al.* (2016) predicted instability index for mouse caltrin I, bovine caltrin and human insulin is 51.99, 24.86 and 43.05, respectively, indicating that while storing in aqueous solution, the mouse caltrin I and human insulin turns in aggregated form with loss of biological activity. Hence, these unstable proteins are not suitable for long term storage in aqueous solution. In our prediction, it is clearly visible that all defence and lytic enzymes are in stable form while storing in aqueous solution.

Another important property is grand averages of hydropathy (GRAVY) which indicate the hydrophilic and soluble behavior of proteins is ranging from  $-0.055$  to  $-0.553$  as this score indicating that all enzymes are hydrophilic in nature except endochitinase 37 as its predicted score is positive which means it is hydrophobic. The aliphatic index defines the relative volume occupied by aliphatic chains of proteins and the increased positive score shows the increased thermostability. Here, all enzymes showed high positive score but alkaline protease is potentially stable over a high temperature range.

### 3.2 Signal peptide prediction

Signal peptide plays important role in targeting the translocation of integral membrane proteins and secretory proteins. For the prediction of signal peptide in all enzymes, SignalP program was used. The SignalP results are summarized in Table 3. SignalP gives output in graphical and numerical values which comprises three different scores, C, S and Y.

**Table 3:** Prediction of signal peptides in defense and lytic enzymes through SignalP

SignalP measure →	Max. C		Max. Y		Mean S		D score		Cleavage site
	Position	Value	Position	Value	Position	Value	Position	Value	
Alkaline proteinase	21	0.820	21	0.875	1-20	0.933	1-20	0.906	between pos. 20 <sup>st</sup> alanine and 21 <sup>st</sup> leucine
Beta-1,3 exo-glucanase	29	0.312	23	0.452	1-22	0.788	1-22	0.634	between pos. 22 <sup>nd</sup> alanine and 23 <sup>rd</sup> threonine
Endo-chitinase 37	26	0.799	26	0.820	1-25	0.847	1-25	0.835	between pos. 25 <sup>st</sup> alanine and 26 <sup>th</sup> glutamine
Endo-chitinase 33	20	0.821	20	0.872	1-19	0.925	1-19	0.901	between pos. 19 <sup>th</sup> alanine and 20 <sup>th</sup> glycine
Endo-chitinase 42	21	0.416	21	0.566	1-20	0.760	1-20	0.671	between pos. 20 <sup>th</sup> serine and 21 <sup>th</sup> serine

The first predicted score is C-score defines “cleavage site” score reported for each position which is significantly high at the cleavage site. For alkaline proteinase enzyme, the reported C-score position indicating first cleavage residue is at 21 position of mature protein which means cleavage site between amino acids 20-21 corresponds to the mature protein starting at position 21 including position 20.

The second predicted score is Y-max which is derivative of C-score combined with S-score resulting better cleavage site prediction rather than raw C-score alone. In a graph, multiple high peaks of C-score is found in one sequence where only one is true cleavage site.

Another predicted score is S-mean and D-score which is for discrimination of secretory and non-secretory proteins. All scores given by SignalP should be very low in case of non-secretory proteins.

### 3.3 Transmembrane region prediction

Transmembrane region prediction is a part of functional study. Predictions were made through SOSUI server which distinguishes

between membrane and soluble protein and predicts membrane helix. The results are tabulated in Table 4. It is clearly visible that alkaline proteinase, beta-1,3-glucan-binding and endochitinase 37 were predicted as membrane proteins and other enzymes, Beta-1,3 exoglucanase, Beta-1,6-glucanase, Catalase-peroxidase, Endochitinase 33 and Endochitinase 42 are soluble proteins. Transmembrane regions are rich in hydrophobic amino acids.

### 3.4 Motif prediction

Identification of motifs aims to get in depth knowledge of molecular and biological activities in uncharacterized proteins. Detection of functionally motifs in proteins of interest allows predicting certain significant features which might even implicate it in certain signalling pathways. Therefore, researches are going on in discovering functionally active motifs in any protein sequence. For domain, motif and functional site analysis PROSITE database was used which catalogued 1810 documentation entries, 1309 patterns, 1218 profiles and 1240 ProRule. The protein sequences of all lytic and

defence enzymes were used as an input for motif search against PROSITE database. Results are tabulated below in the Table 5. The hits were retrieved by matched pattern in case of alkaline proteinase, catalase-peroxidase, endochitinase 33, endochitinase 42, endochitinase 37 and in case of Beta-1,3 exoglucanase, Beta-1,3-

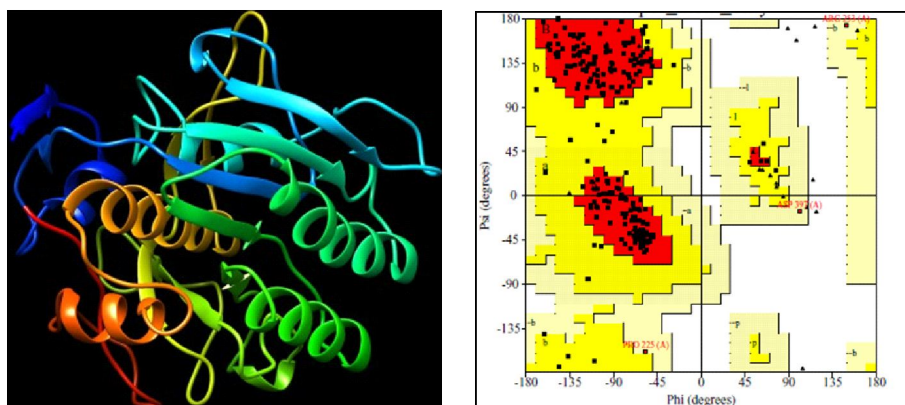
glucan-binding Beta-1,6-glucanase hits were made by matched profile. No hits were found for Beta-1,6 glucanase enzymes. The predicted length of motifs is varied from 10-18 amino acids. It also predicts the conversed active site residues with their behavior.

**Table 4:** Prediction of transmembrane region through SOSUI server

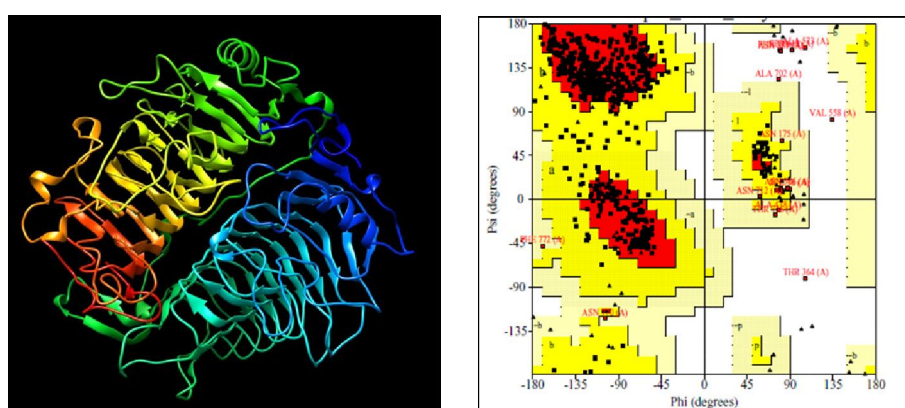
Enzyme name	Transmembrane region	Length	Type of protein
Alkaline proteinase	MAWFKTLALFLFTVTPYVTALPL	23	Membrane protein
Beta-1,3-glucan-binding	TIMPLLGILLGLAISGFLIWDGM	23	Membrane protein
Endochitinase 37	TRLLDASFLLLPVIASTLFGTAS	23	Membrane protein

**Table 5:** Prediction of motifs through PROSITE database

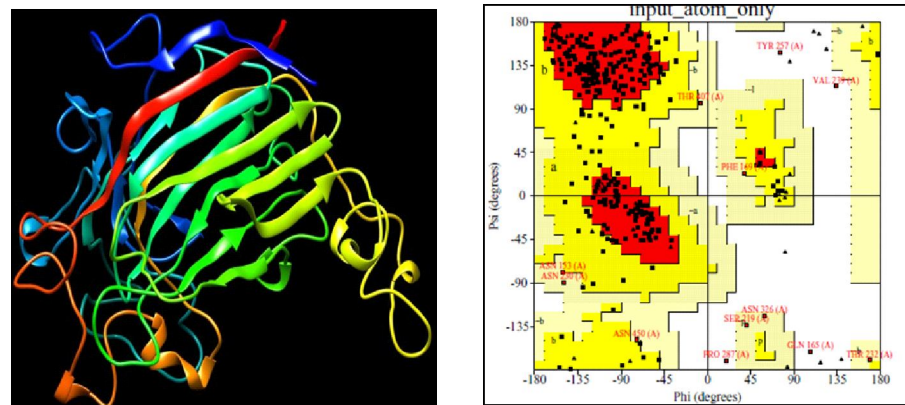
Enzyme names	Description	Position	Matched patterns	Predicted features
Alkaline Proteinase	PS00136 SUBTILASE_ASP Serine proteases, subtilase family, aspartic acid active site	162 - 173	AYVVDSGVftkH	Active site at 166 <sup>th</sup> position
	PS00137 SUBTILASE_HIS Serine proteases, subtilase family, histidine active site	197 - 207	HGThVAGiIAG	Active site at 197 <sup>th</sup> position
	PS00138 SUBTILASE_SER Serine proteases, subtilase family, serine active site	351 - 361	GTSmSaPhVAG	Active site at 353 <sup>th</sup> position
Beta-1,3 exoglucanase	PS51257 PROKAR_LIPOPROTEIN Prokaryotic membrane lipoprotein lipid attachment site profile	1 - 18	MGFIRSAVLSALTFAAAC	Signal peptide from 1 <sup>st</sup> – 17 <sup>th</sup> Lipid at 18 <sup>th</sup> position N-palmitoyl cysteine site Lipid at 18 <sup>th</sup> position S-diacylglycerol cysteine site
	PS51212 WSC domain	816-914	GWNFLGCYSNDvNGRTLANQVQVAGGAsaMSIEACETASESAGYTIAGVEYSGECWC DTKFQNGGGPAsdgSAQCTMTCSGAPQET CGGPNRLDVYSLA	Domain at 816 <sup>th</sup> – 914 <sup>th</sup> position
		927-1024	GWNFRGCYTDSvNARALIAESVPNGP\$M TIEACQSVCKGLGYTLAGLEYADECYCG NSLANGATIAPdgNAGCNMNCAGNAAET CGGPNRLDIYSYG	Domain at 927 <sup>th</sup> - 1024 <sup>th</sup> position
Beta-1,3-glucan-binding protein	PS51762 GH16_2 Glycosyl hydrolases family 16 (GH16)	149 - 458	EDFSNGLDPsiwtkvQVggFGNGeFEqtggdsnvfeNGHlmikatldanlveknvnlldgtCTskdy SCVAATNTTNGNSSVVPPTL----- SGRINTFKGAKIKYGRVEVTAKLPVGDW LWP AIWMLPVNDTYGPWPssGEIDIMESR GNNHTYsqggnniASSALHWGPD PANDAW WKTNNKrkalHTTYSAEFNTFGLEWSQKY LFTYinsrlqvtytnFNKPMWKRgafpdstangtrLV DIWSSSTRDNTPFDFTEFYLLNLAvG GTNG WFEDgmsGKPWLDHSPNAKKDFWNARD TWYPT	Domain at 149 <sup>th</sup> – 458 <sup>th</sup> position GH16 Active site at 293 <sup>rd</sup> position (Nucleophile) Active site at 298 <sup>th</sup> position (Proton donor)
Beta-1,6-glucanase	No Hits			
Catalase-peroxidase	PS00436 PEROXIDASE_2 Peroxidases active site signature	165-176	GPffIRLaWHNA	Active site at 174 <sup>th</sup> position (Proton acceptor)
	PS00435 PEROXIDASE_1 Peroxidases proximal heme-ligand signature	341 - 351	TVALIAGGHAF	
Endo-chitinase 33	PS01095 CHITINASE_18 Chitinases family 18 active site	159 - 167	VDGFDFDfE	Active site at 167 <sup>th</sup> position (Proton donor)
Endo-chitinase 37	PS01095 CHITINASE_18 Chitinases family 18 active site	152 - 160	FDGIDIDiE	Active site at 160 <sup>th</sup> position (Proton donor)
Endo-chitinase 42	PS01095 CHITINASE_18 Chitinases family 18 active site	163 - 171	FDGIDIDwE	Active site at 171 <sup>th</sup> position (Proton donor)



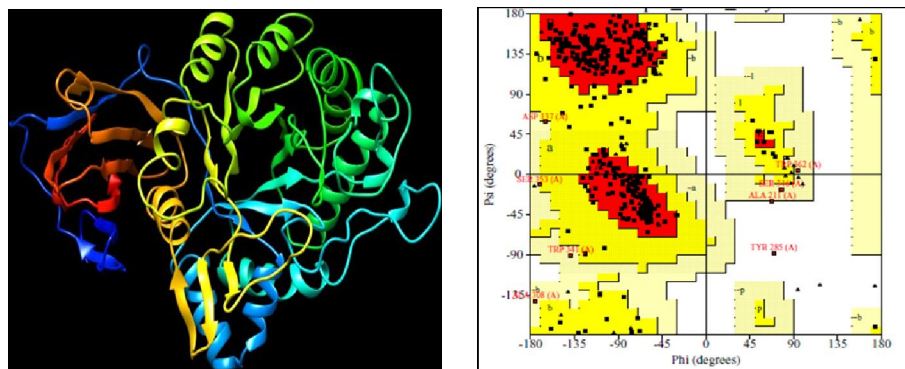
A



B

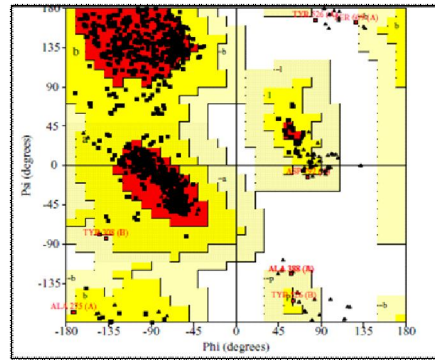
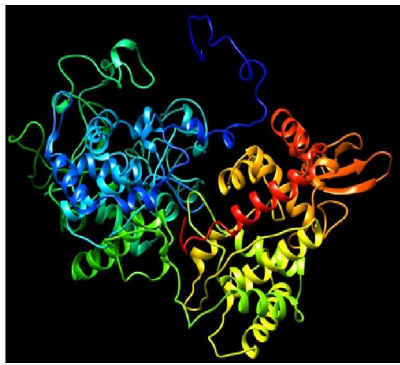


C

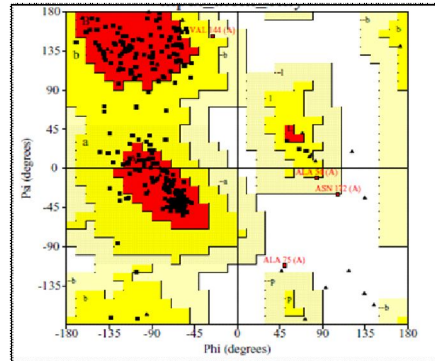
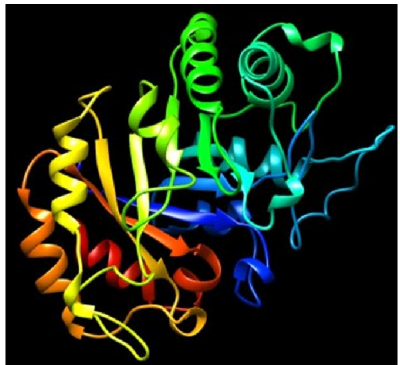


D

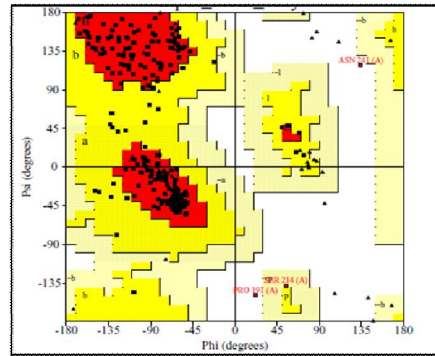




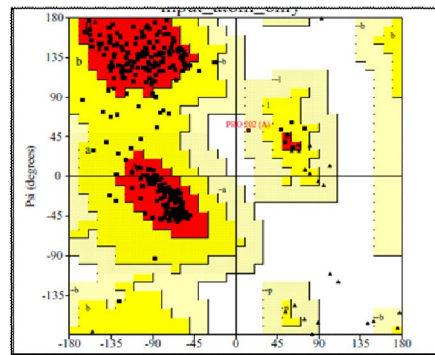
E



F



G



H

**Figure 1:** Homology modelled structure and Ramachandran Plot for (A) Alkaline proteinase, (B) Beta-1,3 exoglucanase, (C) Beta-1,3-glucan-binding protein, (D) Beta-1,6-glucanase, (E) Catalase-peroxidase, (F) Endochitinase 33, (G) Endochitinase 37, and (H) Endochitinase 42.

### 3.5 Template selection, molecular modelling and validation

BLAST search was performed in order to search of crystal structure of homologous proteins against protein data bank (.). Homologues models having sequence identities between 68-25% were chosen as a template structure. Sequence identity percentage (total number of amino acids matched between two different sequences) correlated with coverage percentage (% of query sequence overlaps to template sequence) plays significant role in choosing reliable template structure for comparative modelling. Accuracy of modeled structure depends on when the selected template is having more than 50% of sequence similarity (Baker and Sali, 2001). So, in our work, the highest percentage sequence similarity was obtained for endochitinase 42 of the crystal structure of A chitinase crch1 from the nematophagous fungus, *Clonostachys rosea* (PDB: 3G6L\_A) with 68% sequence similarity and 96% coverage.

In case of beta-1,3 selected template was of PDB ID: 5M5Z\_A with 46% sequence similarity and 72% coverage, followed by endochitinase 33 (PDB: 2UY2\_A with 45% sequence similarity and

89% coverage) and alkaline proteinase (PDB: 3F7O\_A with 42% sequence similarity and 68% coverage). The selected PDB ID 5KZ6\_A (33% sequence similarity and 89% coverage) and 5NGK\_A (31% sequence similarity and 87% coverage) for endochitinase 37 and beta-1,6-glucanase, respectively. The lowest sequence identity (29%) template is of PDB: 3AZX\_A with 57% coverage was found for beta-1,3-glucan-binding protein.

For further analysis, the models were developed by the Swiss-model Workspace. PROCHECK is a Web server, used for structure assessment. It gives Ramachandran Plot which observed residues in favoured, allowed and outer lined regions and it also allows studying steric hindrance. Ramachandran Plot of all enzymes are given below which showed that the residues present in the protein, if found in red colored region denotes that there is no steric hindrance, yellow color denotes some steric hindrance and white color denotes the forbidden zone, where only glycine "G" is found as it has no side chains (Figure 1). It also describes % residues found in favored region and rest of amino acid percentage in outer lined region (Table 6).

**Table 6:** Ramachandran Plot results through PROCHECK

Ramachandran plot $\Rightarrow$	Residues in favoured regions	Residues in allowed regions	Residues in generously allowed regions	Residues in disallowed regions
Alkaline proteinase	88.8%	10.4%	0.8%	0.0%
Beta-1,3 exoglucanase	83.5%	14.2%	1.3%	1.1%
Beta-1,3-glucan-binding protein	78.7%	17.4%	2.8%	1.0%
Beta-1,6-glucanase	83.6%	14.4%	1.5%	0.5%
Catalase-peroxidase	90.9%	8.4%	0.5%	0.2%
Endochitinase 33	82.8%	15.6%	1.2%	0.4%
Endochitinase 37	89.3%	9.8%	0.4%	0.4%
Endochitinase 42	90.5%	9.5%	0.0%	0.0%

### 3.6 Structure ligand binding site identification

Active site prediction is most important aspect as the interaction behavior of proteins with their ligands and other small molecules reveals the cellular function and it also helps to characterized proteins in functional point of view (Singh and Chaube, 2014). Binding site prediction of lytic and defence enzymes of *T. harzianum* were predicted by using SiteHound Web server which identified ligand binding site by computing the interaction energy between protein structure and various chemical probe and carries out cluster analysis. All clusters are, further ranked up according to total interaction energy calculated by molecular interaction field (MIF) (Hernandez *et al.*, 2009). The Web server uses four probes, *i.e.*, methyl probe used to predict binding site for drug like molecules, hydroxyl probe used to characterized sugar binding site, phosphate probe used for the detection of phosphorylated ligands and an aromatic carbon probe used to characterize various other ligands. In a study done by Ghersi and Sanchez (2009), in which test was

conducted on 77 protein-ligand complexes containing drug-like molecule showed 95% correct binding site in case of using methyl probe while using phosphate probe in set 120 protein bound with phosphorylated ligands, it gives similar accuracy results as given in methyl probe.

In case of motif analysis, some features were predicted in case of, beta-1,3-glucan-binding protein, catalase-peroxidase, endochitinase 33, endochitinase 37 and endochitinase 42 which includes active site prediction at different amino acid position. The predicted three active sites of alkaline proteinase at 166, 197 and 353 positions were found interacting with phosphate oxygen probe which suggests that binding of phosphorylated ligand with their identified putative binding site may increase or decrease the catalytic mechanism as proteases penetrates the fungal pathogenic cell wall and leads to lysis.

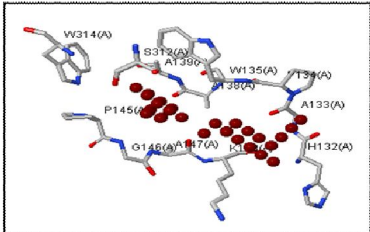
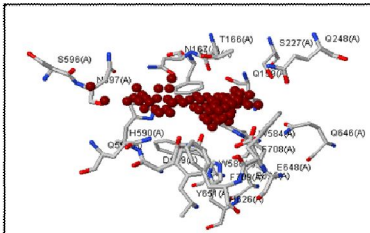
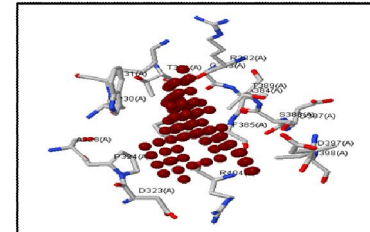
In case of beta-1,3-glucan-binding protein, the predicted active site was found to be interacting with aromatic carbon probe suggesting

that other ligands such as various metal ions and inorganic compound shall increase or decrease the catalytic mechanism of enzyme but it is important to check the other ligands interaction with particular enzymes because in recent studies; it was found that those small ligands causes harmful effect to *T. harzianum* which lead to deactivation of enzymes. A similar *in silico* study was conducted by Gavanji *et al.* (2013), in which they check the possible interaction of chitinase enzyme of *T. harzianum* with silver nitrate ( $\text{AgNO}_3$ ) particles as it has strong antimicrobial activity against pathogenic microbes through docking technique. Their observations suggested that chitinase and silver nitrate have interactions which will lead to

deactivation of enzymes. Another bioinformatics study was conducted by predicting copper, zinc and iron binding proteins with fungal pathogenic species *Paracoccidioides* and the obtained results suggested that these ions played important role in virulence (Tristão *et al.*, 2015).

In case of catalase-peroxidase and endochitinase 37, all predicted active sites were found to be interacting with all four probes whereas endochitinase 33 is interacting with all probes except phosphate oxygen. Endochitinase 42 was found to be interacting with methyl carbon probe. Binding site and descriptors calculations are given in Tables 7-10 for lytic and defence enzymes of *T. harzianum*.

**Table 7:** SiteHound results with methyl carbon (CMET) probe

Enzyme	Energy	Energy range	Volume	Residues in cluster 1	Interaction site
Alkaline proteinase	-589.70	(-17.79, -8.96)	52.00	HIS132A ALA133A PRO134A TRP135A ALA138A ALA139A SER142A LYS144A PRO145A GLY146A ALA147A LYS148A SER312A TRP314A	
Beta-1,3 exo glucanase	-1184.14	(-19.16, -8.93)	98.00	ILE164A THR166A ASN167A PHE169A GLN199A SER227A GLN248A TRP584A TRP586A ASP589A HIS590A SER596A ASN597A GLN599A GLU625A HIS626A GLN646A GLU648A TYR651A TYR706A PHE708A PHE709A ASN715A	
Beta-1,3-glucan-binding protein	-1377.14	(-21.62, -8.94)	112.00	ASP323A PRO324A ALA328A TRP330A LYS331A THR371A ARG382A GLY383A ALA384A PHE385A ASP387A SER388A THR389A ASP397A ILE398A SER400A SER401A ARG404A	





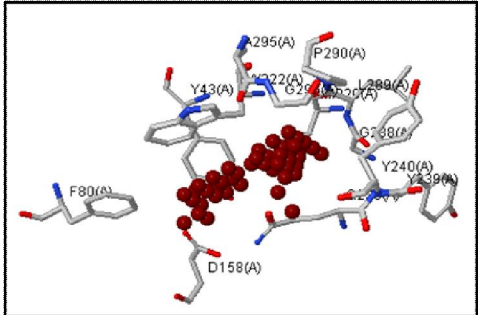
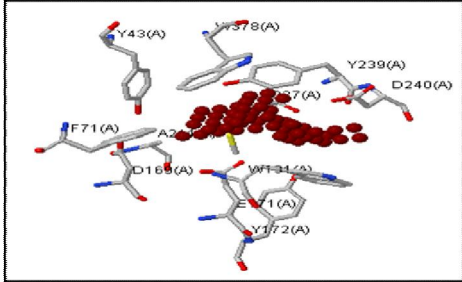
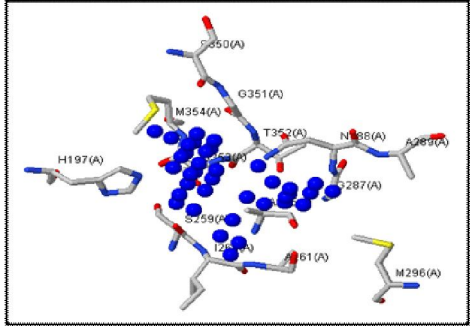
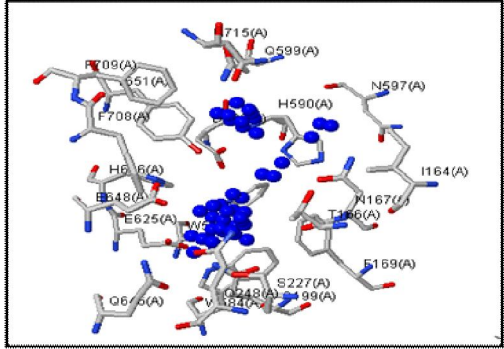
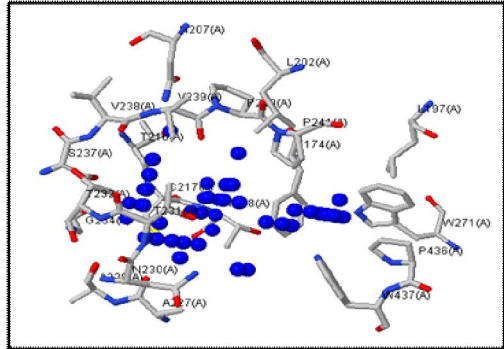
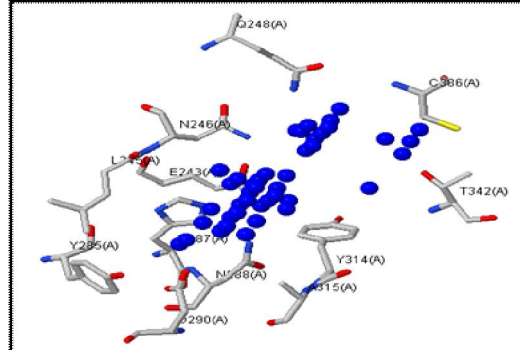
Endo chitinase 37	-939.54	(-21.56, -8.97)	77.00	TYR43A PHE80A GLY123A ASP158A ILE159A <b>GLU160A</b> ALA199A PRO200A GLU201A TYR204A TYR219A GLN238A TYR239A TYR240A ASN241A GLY288A LEU289A PRO290A ALA295A GLY296A GLY297A MET320A TRP322A	
Endo chitinase 42	-967.93	(-20.20, -8.90)	80.00	TYR43A PHE71A TRP131A ASP169A <b>GLU171A</b> TYR172A ALA211A MET237A TYR239A ASP240A TYR293A ARG295A ILE319A TRP378A	

Table 8: SiteHound results with phosphate oxygen (OP) probe

Enzyme ↓	Energy	Energy range	Volume	Residues in cluster 1	Interaction site
Alkaline proteinase	-1913.38	(-29.96, - 8.51)	139.00	<b>ASP166A</b> SER167A GLU195A <b>HIS197A</b> LEU224A SER227A VAL228A SER259A ILE260A ALA261A ALA285A GLY287A ASN288A ALA289A MET296A MET349A SER350A GLY351A THR352A <b>SER353A</b> MET354A	

Beta-1,3 exo glucanase	-1558.93	(-23.33, - 8.53)	120.00	PRO129A ASN175A GLN204A ASN205A ASP233A MET234A THR235A ARG253A ASN254A THR392A ARG393A SER394A LYS395A PRO396A GLN397A GLU399A	
Beta-1,3- glucan- binding protein	-2715.94	(-27.83, - 8.54)	213.00	VAL166A GLY167A GLY168A GLY172A GLU173A PHE174A LEU197A ASP199A LEU202A ASN207A GLY215A THR216A CYS217A THR218A SER219A VAL226A ALA227A ALA228A ASN230A THR231A THR232A GLY234A SER237A VAL238A VAL239A PRO240A PRO241A TRP271A PRO436A TRP437A	
Beta-1,6- glucanase	-2438.03	(-25.40, - 8.50)	191.00	ASP127A TRP180A ASN242A GLU243A ASN246A GLN248A TYR251A HIS287A ASN288A HIS312A TYR314A GLU339A TRP341A THR342A PRO343A THR345A TRP369A GLY384A GLY385A CYS386A GLY387A THR388A	

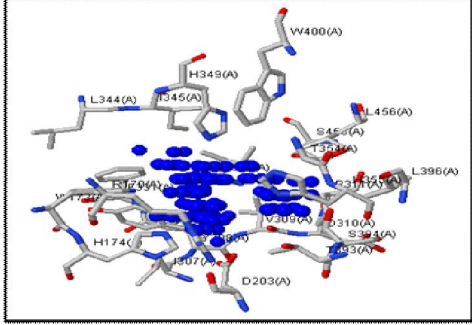
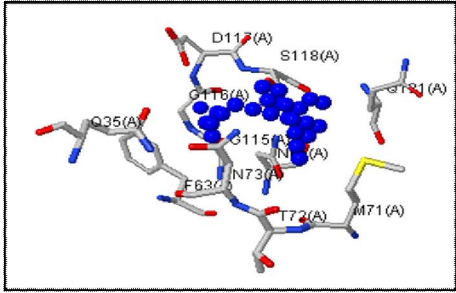
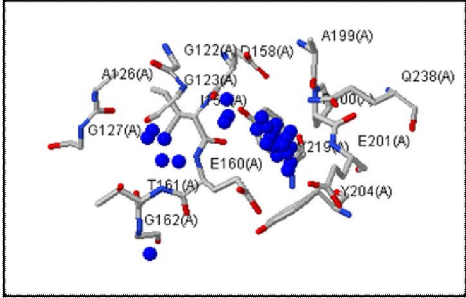
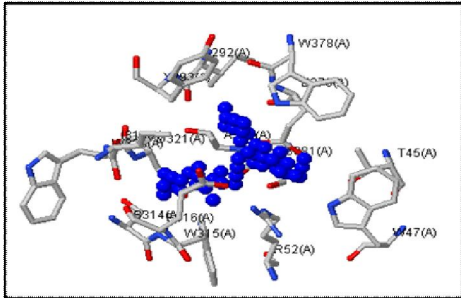
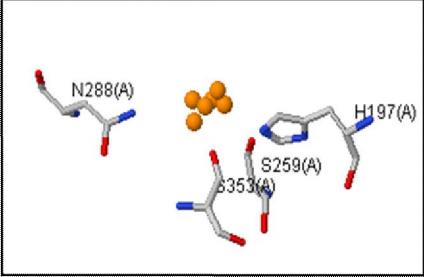
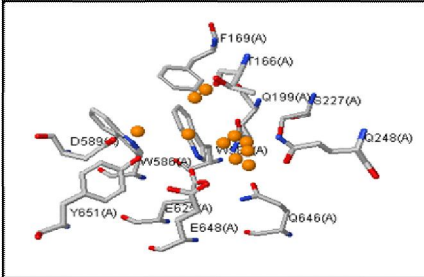
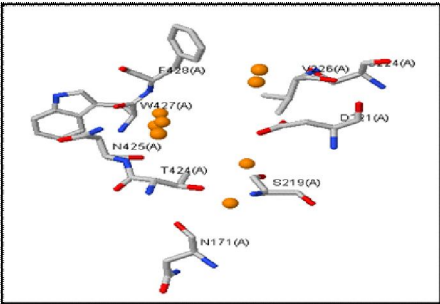
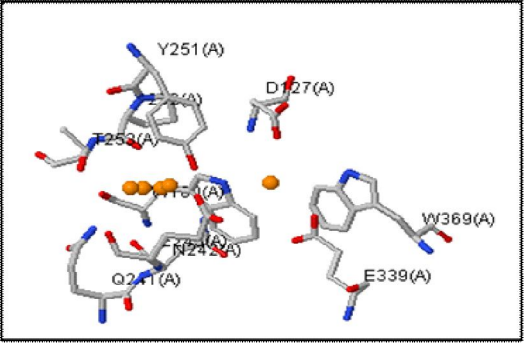
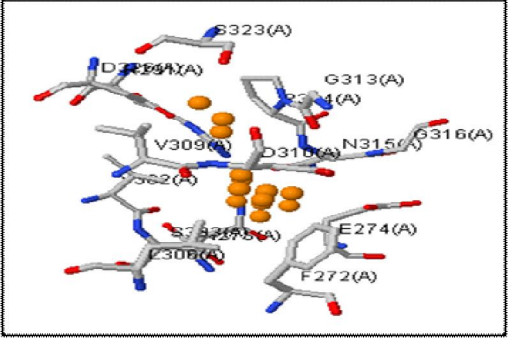
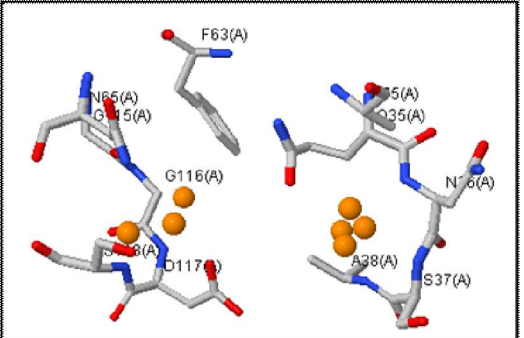
Catalase- peroxidase	-4496.18	(-29.05, - 8.52)	320.00	PRO166A ILE169A ARG170A TRP173A <b>HIS174A</b> ASP203A LEU306A ILE307A TYR308A VAL309A ASP310A PRO311A ILE327A THR330A PHE331A LEU344A ILE345A GLY348A HIS349A PHE351A GLY352A LYS353A THR354A HIS355A THR393A SER394A LEU396A TRP400A LEU456A SER458A ASP459A	
Endo chitinase 33	-2030.74	(-30.18, - 8.54)	145.00	GLN35A ASN36A SER37A ALA38A PHE63A ASN65A GLY66A MET71A THR72A ASN73A ALA75A GLY115A GLY116A ASP117A SER118A GLN121A	
Endo chitinase 37	-1189.43	(-30.24, - 8.57)	84.00	PHE80A ILE83A GLY122A GLY123A ALA124A ALA126A GLY127A ILE128A ILE157A ASP158A ILE159A <b>GLU160A</b> THR161A ALA199A PRO200A GLU201A TYR204A TYR219A GLN238A TRP322A	
Endo chitinase 42	-2398.85	(-34.51, - 8.51)	165.00	THR45A TRP47A GLY48A TYR50A ASP51A ARG52A ILE292A TYR293A SER314A TRP315A GLU316A ILE319A TRP320A ASP321A TRP378A GLU379A SER381A ALA382A	

Table 9: SiteHound results with hydroxyl (OA) probe

Enzyme ↓	Energy	Energy range	Volume	Residues in cluster 1	Interaction site
Alkaline proteinase	-935.78	(-23.34, -8.61)	80.00	HIS197A LEU224A SER259A ILE260A ALA261A SER262A ALA285A GLY287A ASN288A ALA289A MET296A MET349A SER350A GLY351A THR352A SER353A MET354A	
Beta-1,3 exo glucanase	-1119.24	(-20.77, -8.50)	94.00	ILE164A THR166A ASN167A PHE169A GLN199A SER227A GLN248A GLN249A TRP584A TRP586A ASP589A HIS590A ASN597A GLN599A GLU625A HIS626A GLN646A GLU648A TYR651A TYR706A PHE708A PHE709A ASN715A	
Beta-1,3-glucan-binding protein	-1569.66	(-18.71, -8.53)	144.00	VAL166A GLY168A GLU173A PHE174A ASP199A LEU202A ASN207A GLY215A THR216A CYS217A THR218A CYS225A VAL226A ALA227A ALA228A ASN230A THR231A THR232A GLY234A SER237A VAL238A VAL239A PRO240A PRO241A LYS435A TRP437A	



Beta-1,6-glucanase	-941.82	(-18.81, -8.57)	85.00	ASP127A TRP180A ASN242A GLU243A ASN246A GLN248A TYR251A TYR314A GLU339A TRP341A THR342A PRO343A TRP369A GLY384A GLY385A CYS386A THR388A	
Catalase- peroxidase	-2780.23	(-24.15, -8.51)	241.00	PRO166A ILE169A ARG170A TRP173A <b>HIS174A</b> ASP203A ILE307A TYR308A VAL309A ASP310A PRO311A ILE327A THR330A PHE331A LEU344A ILE345A GLY348A HIS349A PHE351A LYS353A THR354A HIS355A THR393A SER394A LEU396A TRP400A LEU456A SER458A ASP459A PHE486A TRP490A	
Endo chitinase 33	-804.44	(-19.05, -8.58)	71.00	TYR32A ALA38A PHE63A GLY116A ASP117A ASP165A <b>GLU167A</b> ALA197A PRO198A GLN199A GLN222A TYR224A ASN225A ALA272A MET299A TRP301A	

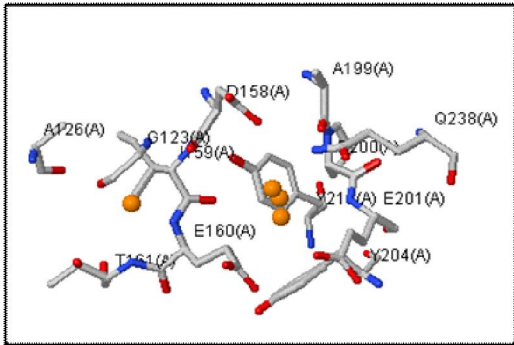
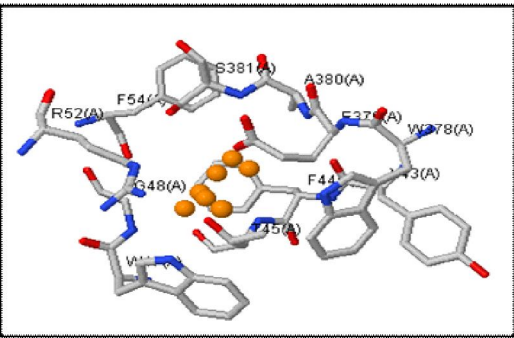
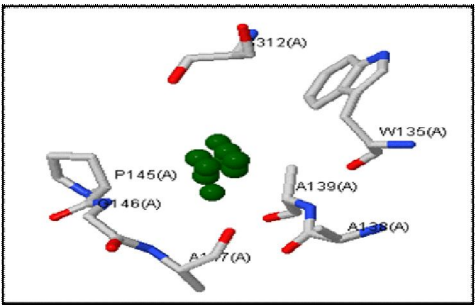
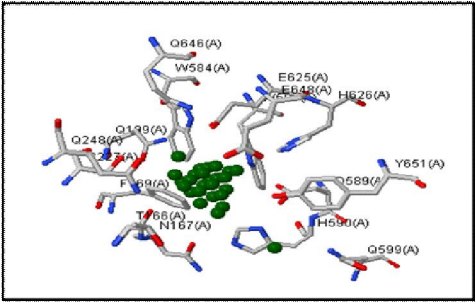
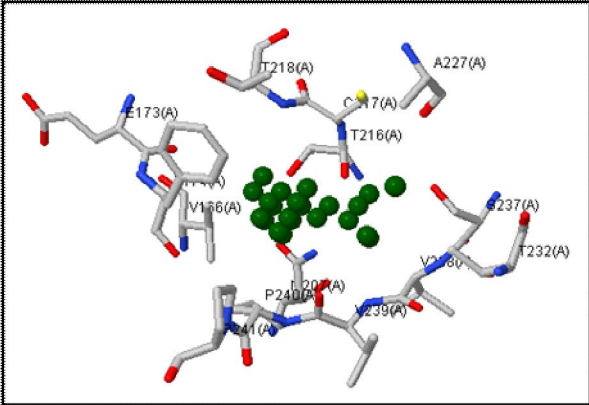
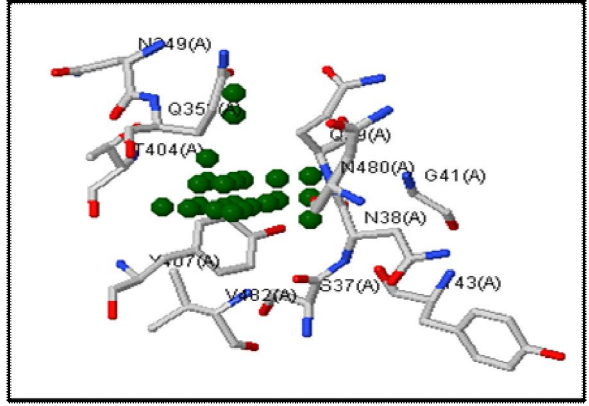
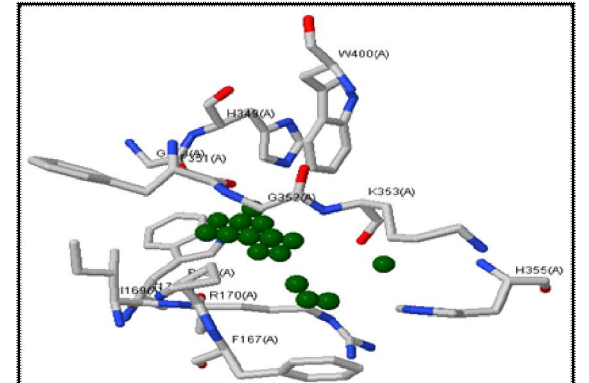
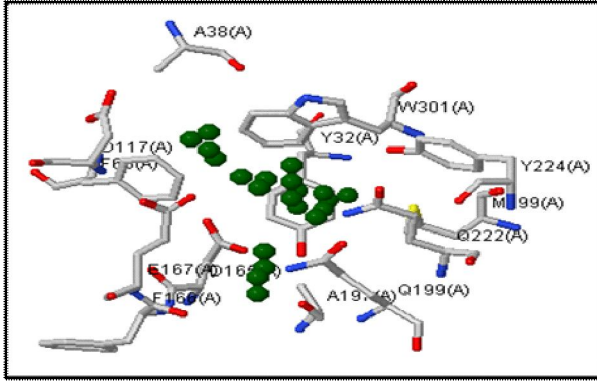
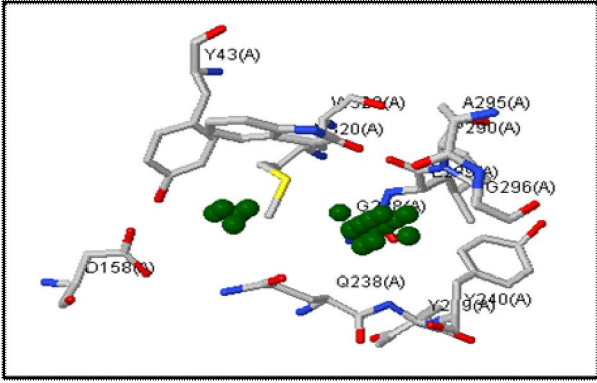
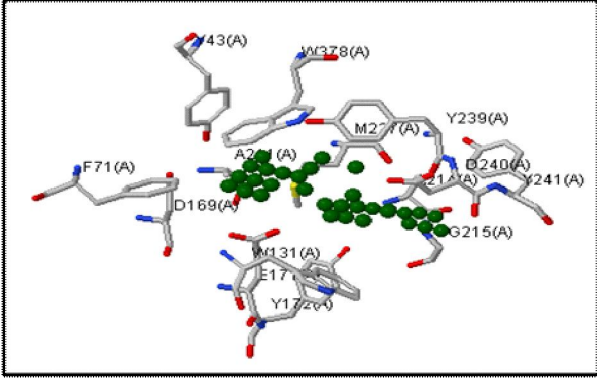
Endo chitinase 37	-614.51	(-18.79, -8.55)	52.00	PHE80A GLY122A GLY123A ALA126A GLY127A ILE157A ASP158A ILE159A <b>GLU160A</b> THR161A ALA199A PRO200A GLU201A TYR204A TYR219A GLN238A	
Endo chitinase 42	-1074.74	(-21.62, -8.58)	90.00	THR45A TRP47A ASP51A ARG52A ASP240A ILE292A TYR293A ARG295A SER314A TRP315A GLU316A ILE319A ASP321A TRP378A GLU379A ALA382A	

Table 10: SiteHound results with aromatic (CR1) probe

Enzyme ↓	Energy	Energy range	Volume	Residues in cluster 1	Interaction site
Alkaline proteinase	-188.04	(-13.64, -8.71)	18.00	HIS132A ALA133A TRP135A ALA138A ALA139A SER142A LYS144A PRO145A GLY146A ALA147A LYS148A SER312A TRP314A	
Beta-1,3 exo glucanase	-448.52	(-16.24, -8.67)	41.00	THR166A ASN167A PHE169A GLN199A SER227A GLN248A TRP584A TRP586A ASP589A HIS590A ASN597A GLN599A GLU625A HIS626A GLU648A TYR651A	

Beta-1,3-glucan-binding protein	-660.34	(-14.57, -8.51)	62.00	PHE169A GLY170A ASN171A GLU173A ARG246A TRP273A ALA275A TRP277A LEU279A TRP288A <b>GLU293A</b> ASP295A <b>GLU298A</b> HIS319A ASN326A TRP329A ASN418A ALA420A	
Beta-1,6-glucanase	-500.58	(-13.27, -8.55)	49.00	SER37A ASN38A GLN39A GLY41A TYR43A TRP348A ASN349A GLN350A PHE402A GLN403A THR404A TYR407A ASN480A VAL482A	
Catalase- peroxidase	-506.21	(-14.59, -8.51)	50.00	TRP173A <b>HIS174A</b> ASP203A ILE307A TYR308A VAL309A LEU344A ILE345A GLY348A HIS349A SER458A ASP459A TRP490A	

Endo chitinase 33	-398.27	(-14.01, -8.50)	38.00	TYR32A ALA38A SER40A PHE63A ASP117A ASP165A <b>GLU167A</b> GLN222A TYR224A ALA272A MET299A TRP301A	
Endo chitinase 37	-394.20	(-16.66, -8.56)	37.00	TYR43A PHE80A GLY123A ASP158A ILE159A <b>GLU160A</b> ALA199A PRO200A GLN238A TYR239A TYR240A GLY288A LEU289A PRO290A ALA295A MET320A TRP322A	
Endo chitinase 42	-430.47	(-13.64, -8.59)	42.00	TYR172A PRO213A ALA214A GLY215A ASN218A MET237A TYR239A ASP240A TYR241A ALA242A GLY243A PHE245A SER246A PHE267A ARG295A	

The aim of using SiteHound Web server is to identify different ligand binding sites with the use of four different probes which corresponds to putative binding sites. The use of different chemical probes allows to study and characterize different potential binding site which helps to understand the fundamental mechanism of protein ligand interaction.

#### 4. Conclusion

In *Trichoderma*, mycoparasitism is one of major mechanism involved in their antagonistic activity against phytopathogens. Multiple enzymes play crucial role in mycoparasitic interaction. In this study, lytic and defense enzymes were selected for *in silico* studies to improve their potential functional aspects. Due to lack of experimental structures of lytic and defence enzyme, *in silico* structural analysis was performed using homology modelling approach and further validated. Prediction of motifs in protein sequences allows studying conserved patterns throughout the evolutionary process. *In silico* identification and prediction of potential of binding site is very crucial feature of structure-based drug designing. But, here in this study, prediction of different ligand binding sites, using chemical probes will provide to study interaction between ligands and enzymes to improve the efficacy of biological experiment for the development of small compounds which increases the catalytic mechanism of enzymes, involves in mycoparasitism and will lead to destruction of phytopathogens.

#### Acknowledgements

The authors are grateful for the Financial Support given by the Council of Science and Technology, Lucknow.

#### Conflict of interest

We declare that we have no conflict of interest.

#### References

- Arnold, K.; Bordoli, L.; Kopp, J. and Schwede, T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*, **22**:195-201.
- Baker, D. and Sali, A. (2001). Protein structure prediction and structural genomics. *Science*, **294**:93-96.
- Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, L.; Shindyalov, I. and Bourne P. E. (2000). The Protein Data Bank. *Nucleic Acids Res.*, **28**:235-242.
- Bindu, V. B. B.; Srinath, M.; Shailaja, A. and Giri, C. C. (2017). Comparative protein profile studies and *in silico* structural/functional analysis of HMGR (ApHMGR) in *Andrographis paniculata* (Burm.f.) Wall. ex Nees. *Ann. Phytomed.*, **6**(1):30-44.
- Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N., Papadopoulos, J.; Bealer, K. and Madden, T. L. (2009). BLAST: Architecture and applications. *BMC Bioinformatics*, **10**:421-430.
- Chakraborty, M. R. and Chatterjee, N. C. (2007). Interaction of *Trichoderma harzianum* with *Fusarium solani* during its pathogenesis and the associated resistance of the host. *Asian J. Exp. Sci.*, **21**:353-357.
- De Castro, E.; Sigrist, C. J. A.; Gattiker, A.; Bulliard, V.; Langendijk Genevaux, P. S.; Gasteiger, E.; Bairoch, A. and Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res.*, **1**: 34 (Web Server issue): W362-5.
- Devi, N. and Azmi, W. (2018). Structural analysis and characterization of a clinically important low molecular weight natural dextran synthesized by *Leuconostoc lactis* KU665298 dextranase. *Ann. Phytomed.*, **7**(1):52-62.
- Fotoohiyan, Z.; Rezaee, S.; Shahidi Bonjar, H. Gh.; Mohammadi, H. A. and Moradi, M. (2015). Induction of systemic resistance by *Trichoderma harzianum* isolates in Pistachio plants infected with *Verticillium dahlia*. *Journal of Nuts*, **6**(2):95-111.
- Gavanji, S.; Aziz, A. H.; Larki, B. and Mojiri, A. (2013). Computational prediction and analysis of interaction of silver nitrate with chitinase enzyme. *IJSRES*, **1**(4):50-62.
- Gherzi, D. and Sanchez, R. (2009). Easy MIFS and Site Hound: A toolkit for the identification of ligand-binding sites in protein structures. *Bioinformatics*, **25**:3185-3186.
- Grasso, J. E.; Sottile, E. A. and Coronel, E. C. (2016). Structural prediction and *in silico* physicochemical characterization for mouse caltrin I and bovine caltrin proteins. *Bioinformatics and Biology Insights*. **10**:225-236 doi: 10.4137/BBI.S38191.
- Green, H.; Larsen, J.; Olsson, A. P.; Jensen, F. D. and Jakobsen, I. (1999). Suppression of the biocontrol agent *Trichoderma harzianum* by mycelium of the arbuscular mycorrhizal fungus *Glomus intraradices* in root-free soil. *Applied and Environmental Microbiology*, **65**(4):1428-1434.
- Haran, S.; Schickler, H. and Chet, I. (1996). Molecular mechanisms of lytic enzymes involved in the biocontrol activity of *Trichoderma harzianum*. *Microbiology*, **142**:2321-2331.
- Hernandez, M., Gherzi, D. and Sanchez, R. (2009). SITEHOUND-web: A server for ligand binding site identification in protein structures. *Nucleic Acids Research*, **37**:W413-W416 doi:10.1093/nar/gkp281.
- Hirokawa, T.; Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**(4):378-379.
- Khaled, L. B.; Pérez-Gilbert, M.; Dreyer, B.; Oihabi, A.; Honrubia, M. and Morte, A. (2008). Peroxidase changes in *Phoenix dactylifera* palms inoculated with mycorrhizal and biocontrol fungi. *Agron. Sustain. Dev.*, **28**(3):411-418.
- Laskowski, R. A.; MacArthur, M. W.; Moss, D. S. and Thornton J. M. (1993). PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.*, **26**:283-291.
- Petersen, N. T.; Brunak, S.; Heijne, V. G. and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*: doi:10.1038/nmeth.1701.
- Shi, W. and Chance, M. R. (2008). Metallomics and metalloproteomics. *Cell.Mol. LifeSci.*, **65**:3040-3048. doi:10.1007/s00018-008-8189-9.
- Singh, A. and Chaube, R. (2014). Bioinformatic analysis, structure modeling and active site prediction of aquaporin protein from catfish heteropneustes fossilis. *IJRITCC*, **2**(10):3208-3215.
- Strakowska, J.; Błaszczak, L. and Chelkowski, J. (2014). The significance of cellulolytic enzymes produced by *Trichoderma* in the opportunistic lifestyle of this fungus. *J. Basic Microb.*, **54**:S2-S13.
- The UniProt Consortium (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Res.*, **45**:D158-D169.
- Tristão, G. B.; Assunção, Ldo. P.; Dos Santos, L. P.; Borges, C. L.; Silva-Bailão, M. G.; Soares, C. M.; Cavallaro, G. and Bailão, A. M. (2015). Predicting copper, iron, and zinc-binding proteins in pathogenic species of the Paracoccidioides genus. *Front. Microbiol.*, **9**(5):761. doi: 10.3389/fmicb.2014.00761.
- Ward, M. A. D.; Georgopoulos, G.; Hollomon, D. W.; Ishii, H.; Leroux, P.; Ragsdale, N. N. and Schwinn, F. J. (1993). Chemical control of plant diseases: Problems and prospects. *Ann. Rev. Phytopathol.*, **34**:403-421.
- Wilkins, M. R.; Gasteiger, E. and Bairoch, A. (1999). Protein identification and analysis tools in the ExpASY server. *Methods Mol Biol.*, **112**:531-552.