

A comparative proteomic approach to analyse structure, function and evolution of rice chitinases: a step towards increasing plant fungal resistance

Kishore Sarma · Budheswar Dehury · Jagajjit Sahu ·
Ranjan Sarmah · Smita Sahoo · Mousumi Sahu ·
Priyabrata Sen · Mahendra Kumar Modi ·
Madhumita Barooah

Received: 9 February 2012 / Accepted: 14 May 2012 / Published online: 9 June 2012
© Springer-Verlag 2012

Abstract Glycoside hydrolase family 19 chitinases (EC 3.2.1.14) widely distributed in plants, bacteria and viruses catalyse the hydrolysis of chitin and play a major role in plant defense mechanisms and development. Rice possesses several classes of chitinase, out of which a single structure of class I has been reported in PDB to date. In the present study an attempt was made to gain more insight into the structure, function and evolution of class I, II and IV chitinases of GH family 19 from rice. The three-dimensional structures of chitinases were modelled and validated based on available X-ray crystal structures. The structural study revealed that they are highly α -helical and bilobed in nature. These enzymes are single or multi domain and multi-functional in which chitin-binding domain (CBD) and catalytic domain (CatD) are present in class I and IV whereas class II lacks CBD. The CatD possesses a catalytic triad which is thought to be involved in catalytic process. Loop III, which is common in all three classes of chitinases, reflects that it may play a significant role in their function. Our study also confirms that the absence and presence of different loops in GH family 19 of rice may be responsible for various sized products. Molecular phylogeny revealed chitinases in monocotyledons and dicotyledons differed from each other forming two different clusters and may have evolved differentially. More structural study of this enzyme from different plants is required to enhance the knowledge of catalytic mechanism and substrate binding.

Keywords Chitinase · Comparative modeling · Phylogeny · Proteomic

Introduction

Plants represent the major component of earth's biota and are capable of producing their food through the process of photosynthesis. A multitude of pathogens cause loss in the crop yield every year which is often severe across the globe. Plants are equipped with a variety of defense mechanisms to protect themselves against the attack of pathogens. Some of these are constitutive while others are induced upon the attack by pathogens. The interaction of plant and pathogens induces a variety of defense mechanisms which include cell wall strengthening [1], de novo production of antimicrobial compounds (pathogenesis response proteins) and secondary metabolites [2, 3]. In the case of pathogenesis related proteins (PR), chitinase and glucanase play a crucial role since they attack directly on the fungal and insect structural component whereas enzymes of plant secondary metabolite pathway including Chalcone synthase [4] and Phenylalanine ammonia lyase [5] play significant roles due to antimicrobial nature of secondary metabolites.

Chitinases (EC 3.2.1.14), which are found in a wide range of organisms catalyse the hydrolysis of chitin and play a major role in plant defense mechanism against fungal pathogens. Chitinase catalyses the hydrolysis of β -1-4-linkage of the N-acetylglucosamine polymer of chitins, a major component of fungal cell walls [6]. Plant chitinases usually have a wide range of optimum pH (pH 4–9) and are generally stable at temperature up to 60 °C [7]. These enzymes

K. Sarma · B. Dehury · J. Sahu · R. Sarmah · S. Sahoo · M. Sahu ·
P. Sen · M. K. Modi · M. Barooah (✉)
Agri-Bioinformatics Promotion Programme, Department of
Agricultural Biotechnology, Assam Agricultural University,
Jorhat 785013 Assam, India
e-mail: m17barooah@yahoo.co.in

usually have a molecular weight ranging from 25 to 35 kDa. They are usually involved with both active and passive defense against pathogens [8–10]. These enzymes also regulate growth and development by generating or degrading signal molecules [11–13] through programmed cell death (PCD) [14, 15]. Production of chitinases is regulated by a variety of stress factors, both biotic and abiotic, including infection, wound, drought, cold, ozone, heavy metals, excessive salinity and UV light [7, 9, 16–18]. In addition phytohormones, such as ethylene, jasmonic acid, salicylic acid, auxin and cytokinin, induce chitinase expression [19].

Several workers have proposed different classification schemes of plant chitinases and the classification scheme has undergone several modifications. Based on the hydrolytic sites, plant chitinases are broadly classified into two categories, endochitinases and exochitinases. With regard to physicochemical properties and enzymatic activity plant chitinases are classified as PR proteins [7]. Based on biological properties, enzyme activity and coding sequence similarities, chitinases are represented by four PR proteins designated as PR-3, PR-4, PR-8, and PR-11 [20]. Based on the amino acid sequence homology, structure, substrate specificity, mechanisms of catalysis and sensitivity to inhibitors they are classified into seven classes, i.e. I–VII [21]. Class I is further divided into two sub classes, Ia and Ib. PR-3 includes chitinases of class Ia, Ib, II, IV, VI and VII, chitinases of class III belong to PR-8, and chitinases of class V to PR-11. Among the chitin binding proteins, protein with low endochitinase activity form the PR-4 class [22].

Class I chitinases have a cysteine-rich chitin-binding domain (CBD) at N-terminal and a C-terminal catalytic domain (CatD). The CBD is linked to the CatD by a proline and glycine rich linker, which varies in its length and composition [20]. In contrast class II chitinases lack the N-terminal CBD domain and linker region but they show high sequence homology with CatD of class I chitinases. Class III chitinases show lysozyme activity and do not reveal any sequence homology to either class I or II chitinases. All plant chitinases from this class show good percentage of sequence similarity among themselves but differ widely in their isoelectric point [23]. Class IV chitinases have low sequence similarity with class I chitinases; they contain a CBD and CatD which resembles that of class I chitinases. Both domains of class IV chitinases are significantly smaller than those of class I because of one deletion in the CBD and three deletions in CatD. Class V chitinases possess two CBDs in tandem [24, 25]. Heavily truncated CBD along with a proline rich spacer is dominant in class VI chitinases [26]. Class VII chitinases lack the CBD but possess CatD, which is homologous to Class IV chitinases [27].

Furthermore, chitinases are classified into two families of glycoside hydrolases (GH), families 18 and 19, in which glycoside hydrolases are divided into more than 110 families based on the amino acid sequence similarity of their catalytic domains [28, 29]. The members of two different families differ in their amino acid sequences, three-dimensional (3D) structures, and molecular mechanisms of catalytic reactions [30], and are thus considered to have different evolutionary origins. GH family 18 chitinases [class III and V] are widely distributed in a variety of organisms, such as bacteria, fungi, viruses, animals and higher plants. The distribution of GH family 19 [class I, II, IV, VI, VII] enzymes is more restricted and they are mainly found in higher plants and some bacteria [28, 31].

Rice (*Oryza sativa* L.) the model plant represents the monocotyledons whose genome was completely sequenced in 2004. Genome annotation study of rice reveals that chitinase genes are present in all chromosomes except chromosome number seven [32]. Rice possesses several classes of chitinases encoded by different genes located in different chromosomes. Although the expression of these genes is differentially induced and regulated, they act both directly and indirectly in plant defenses as well as are associated with numerous roles in plant physiological function. Several lines of evidence reflect that chitinase expression and enzyme activity has a major contribution in disease resistance against fungal pathogens in rice. Transgenic plants which constitutively expressed a rice class-I chitinase gene, Cht-2 or Cht-3, showed significant resistance against two races of *Magnaporthe grisea* [33]. The activity of chitinase in the transgenic plants, overexpressing a rice chitinase gene, was found to be correlated to the levels of enhanced resistance against *Rhizoctonia solani* [34]. The study on anti-fungal properties of class I and class II chitinases showed that class-I have three to five times higher activity than that of class-II [28]. At pH 3–5 it was shown that purified rice basic class III chitinase, expressed in *Pichia pastoris*, is an effective lytic agent of *Micrococcus lysodeikticus*, but shows a weak fungal inhibition towards *Trichoderma reesei* [35]. Study on class I and class IV chitinases in rice showed that both possess a similar catalytic domain and had similar N-acetyl-chitin-oligosaccharide degradation efficiency [27]. These evidences confirm the hypothesis that rice chitinases play an important role in fungal resistance.

Six crystal structures are available for GH-19 chitinases from plant origin in protein databank (PDB). These include barley (class-II), jack-bean (class-II), mustard (class-I), papaya (class-II), Norway spruce (class-IV) and rice (class-I). Rice possesses several family 19 chitinases. To date only the rice class I chitinase *OsChia1b*, also referred to as *RCC2* or *Cht-2*, which has been reported in PDB [PDB accession code: 2DKV] [28]. The reports on 2DKV reveals that this chitinase is comprised of two domains (N-terminus CBD

Table 1 Rice chitinase sequences analysed in this study

Accession number (UniProtKB)	Protein name	Length of the protein	Classification of rice chitinases		
			Family	Sub-family	Pathogenesis related protein (PR) class
Q42993	Chitinase1	323	Glycosyl hydrolase 19	class I	PR3
Q7DNA1	Chitinase2	340	Glycosyl hydrolase 19	class I	PR3
P24626	Chitinase3	320	Glycosyl hydrolase 19	class I	PR3
O04138	Chitinase4	229	Glycosyl hydrolase 19	class IV	PR3
Q7Y1Z0	Chitinase5	288	Glycosyl hydrolase 19	class IV	PR3
Q6K8R2	Chitinase6	271	Glycosyl hydrolase 19	class IV	PR3
Q7Y1Z1	Chitinase7	340	Glycosyl hydrolase 19	class I	PR3
Q7XCK6	Chitinase8	261	Glycosyl hydrolase 19	class II	PR3
Q688M5	Chitinase9	334	Glycosyl hydrolase 19	class I	PR3
Q5NB11	Chitinase10	290	Glycosyl hydrolase 19	class I	PR3
Q10S66	Chitinase11	256	Glycosyl hydrolase 19	class II	PR3
P25765	Chitinase12	326	Glycosyl hydrolase 19	class I	PR3

and C-terminus CatD) which are interconnected by a linker peptide rich in proline and threonine amino acids. Kezuka et al. also reported that the CBD of 2DKV binds to chitin which acts as an anchor whereas CatD degrades the chitin chain depending upon the linker length. As rice possesses several classes of chitinases, in the present study we have selected a few reviewed representative chitinases of classes

I, II and IV. An attempt was made to elucidate the structure, function and evolution of nine rice (*Oryza sativa* L.) chitinases belonging to family-19 using comparative proteomic approach with the aid of high-throughput computational tools. This study will provide an insight into the structural variations, evolution and molecular function of different classes of chitinase in rice.

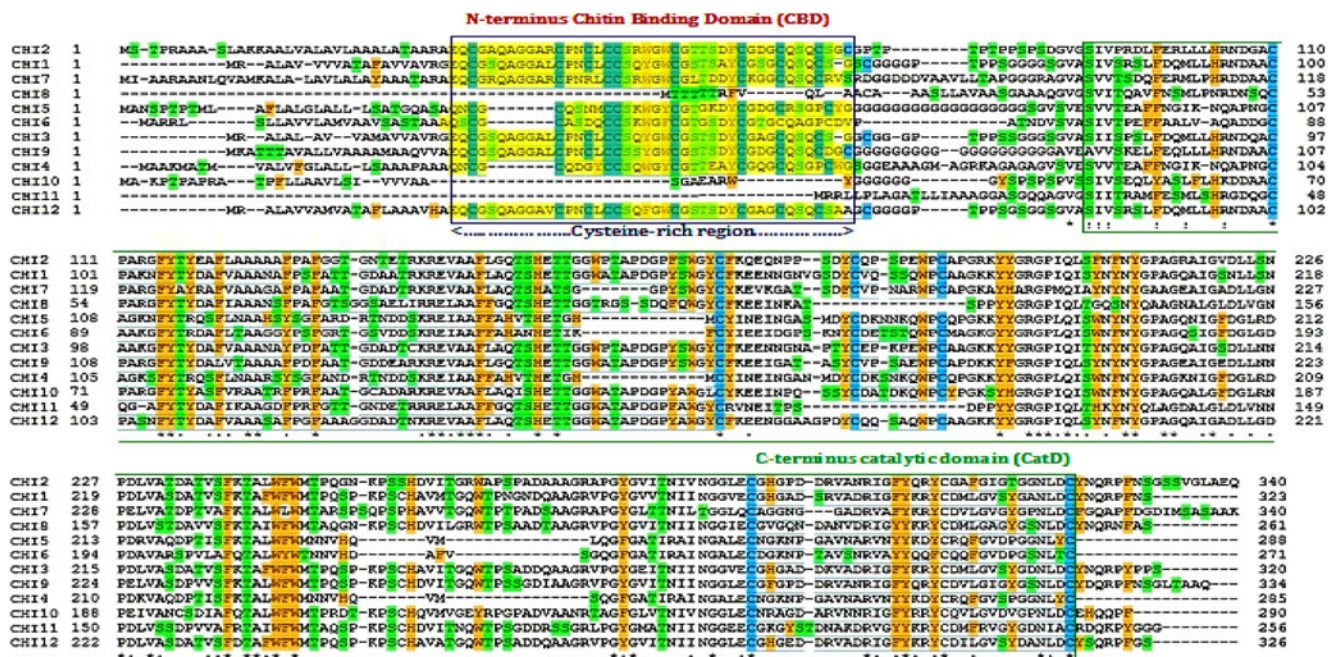


Fig. 1 Amino acid sequence alignment class I (1,2,3,7,9,10,12), class II (8,11) and class IV (4,5,6) chitinases of rice. Sequence alignment was performed in ClustalW. Aromatic amino acid residues conserved

in all sequences are indicated by orange, cysteine residues in cyan, serine and threonine are coloured in green. The two domains (CBD and CatD) are highlighted in square boxes

Methods and materials

Software used

- MEGA5.0
- Modeller9v9

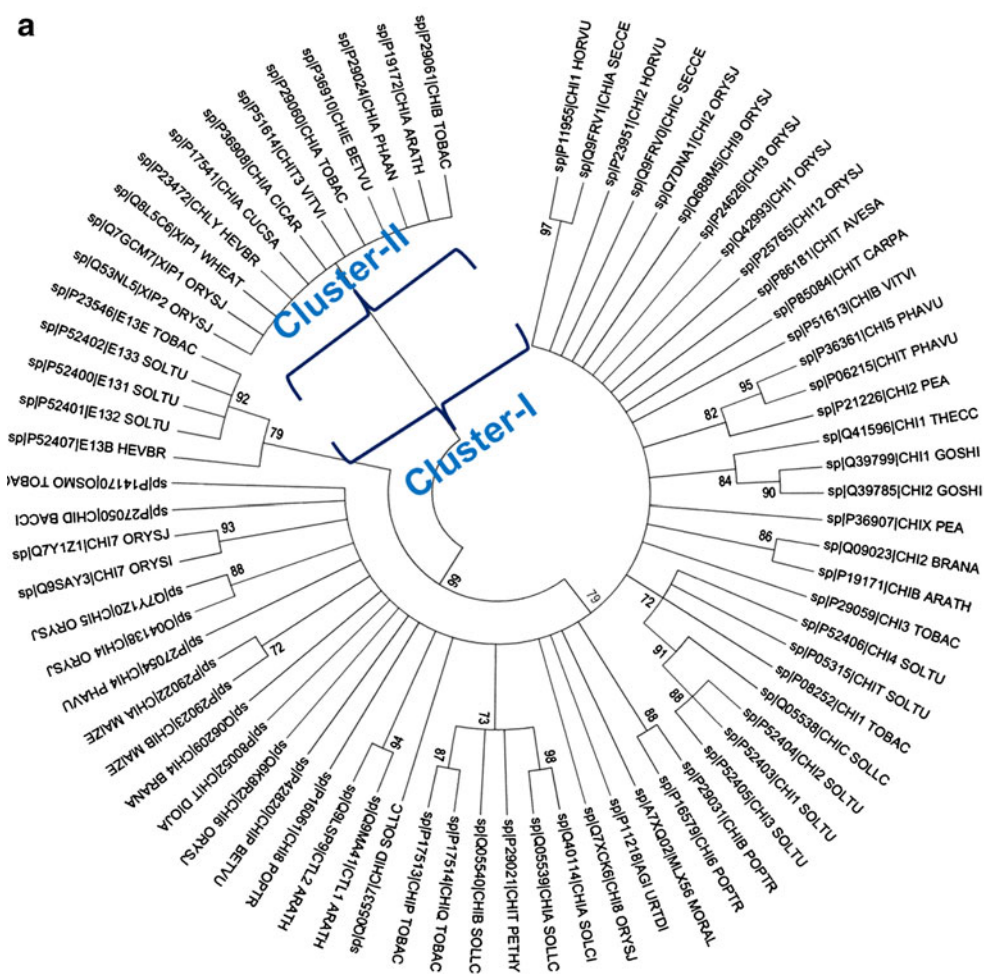
Sequence retrieval and multiple sequence alignment

The reviewed fasta sequences of chitinase enzyme (EC 3.2.1.14) of rice (*Oryza sativa* L.) were retrieved from the UniProtKB (www.uniprot.org/help/uniprotkb) database. In this study a total of 12 chitinase sequences were selected and nine of them were analysed. Information about the twelve sequences is depicted in Table 1. The multiple sequence alignment was produced using ClustalW2 at EMBL-EBI (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) [36].

Phylogeny analysis

For the construction of molecular evolutionary genetic tree, a total of 74 different plant chitinase sequences (reviewed) were

Fig. 2 (a) Circular tree showing the evolutionary origin and evolutionary relationship among different plant chitinases. (b) Phylogenetic tree inferred from neighbour-joining method showing dichotomy with two different clusters with their sub-clusters



retrieved from UniprotKB database and were aligned in ClustalW. Neighbour-joining method [37] was used for the construction of phylogenetic tree using MEGA5.0 [38]. The level of confidence was estimated using bootstrap of 1000 replications.

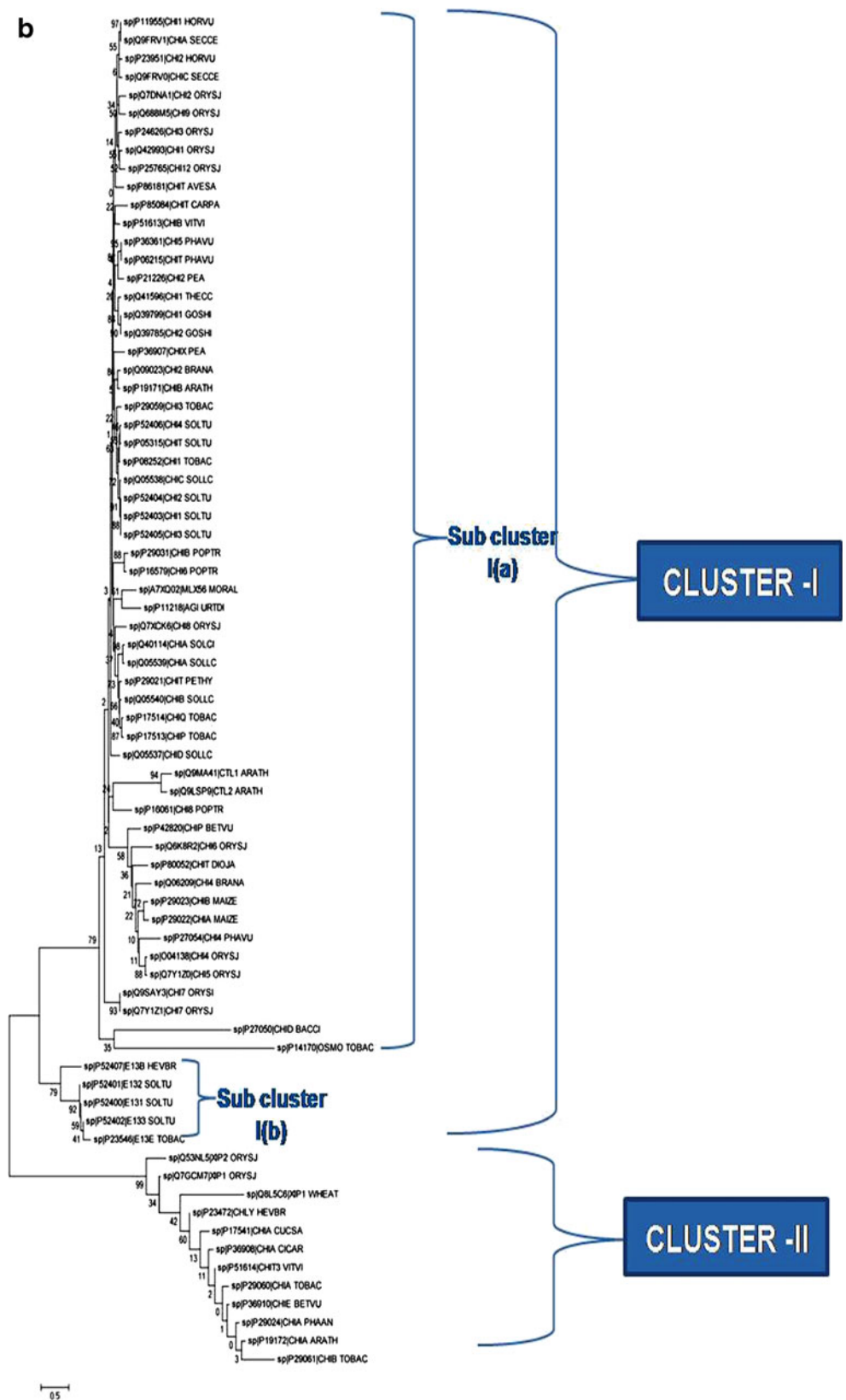
Physico-chemical property analysis of chitinases

For the elucidation of physico-chemical properties of chitinases, ProtParam tool (<http://expasy.org/cgi-bin/protparam>) [39] of ExPASy Proteomic Server was used. The theoretical isoelectric point (pI), molecular weight, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) was calculated.

Secondary structure and disorder region prediction

The secondary structures of chitinases were predicted using PSIPRED server (<http://bioinf.cs.ucl.ac.uk/psipred/>) [40]. This is a simple and accurate secondary structure prediction method that incorporates two feed-forward neural networks which perform an analysis on output obtained from PSI-BLAST. SOPMA (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_

Fig. 2 (continued)



automat.pl?page=npsa_sopma.html) [41] and GOR IV (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_

gor4.html) [42] servers were also used to predict the secondary structures. The disorder regions of rice chitinases were

Table 2 Physico-chemical Properties of chitinases

Protein	Molecular weight (Da)	Theoretical pI	Instability index	Aliphatic index	GRAVY
Chitinase1	33764.5	7.37	35.89	54.77	-0.182
Chitinase3	33681.4	4.84	33.43	53.50	-0.292
Chitinase5	30486.9	8.31	30.37	50.90	-0.419
Chitinase7	35299.7	8.32	37.09	65.15	-0.081
Chitinase8	27551.6	6.09	31.72	67.82	-0.118
Chitinase9	34401.1	4.48	39.72	62.07	-0.121
Chitinase10	31264.1	8.01	46.61	64.72	-0.229
Chitinase11	27747.9	6.42	38.52	58.05	-0.453
Chitinase12	33636.1	4.63	37.53	56.20	-0.107

predicted by protein disorder meta-prediction server (metaPrDOS) (<http://prdos.hgc.jp/meta/>) [43].

Domain analysis and linker prediction

For characterisation and understanding of protein function, detailed knowledge of protein domain boundaries and architecture is essential. In the absence of known 3D structures the delineation of domain boundaries of a given sequence benefits many areas in protein science, such as protein engineering and protein structure prediction [44]. We used comparative domain boundary prediction method called SBASE (<http://hydra.icgeb.trieste.it/sbase/>) [45] for domain boundary prediction. This method exhaustively searches the sequence against known domain definitions within the associated domain database. Domain boundaries along with domain contents are predicted and thus this can be useful for the identification of protein domain architecture. Linkers are the sequence region between defined structural domains. Linker regions are usually non-globular, un-

structured or a low complexity segment that is flexible in 3D-space, but studies showed that linker region may significantly affect the cooperation and interaction between domains and therefore alter the overall functionality and efficiency of multiple domain proteins [46]. The linker sequences joining the two discrete domains of chitinases were delineated manually.

Comparative modelling

Template identification

In order to find suitable templates for the comparative modelling of chitinase, the target sequences of chitinases were searched for similar sequences using the basic local alignment search tool (BLASTP) [47] against PDB. The Blosum-62 matrix was used with a default threshold E-value of 10 and inclusion threshold value of 0.005. Templates were selected based on the query coverage, sequence identity, less E-value and structural resolution.

Table 3 Secondary structures statistics predicted from different servers

Chitinase	GORIV			SOPMA			PSIRED		
	Helix (in %)	Random coil (in %)	Extended strand & beta turn (in %)	Helix (in %)	Random coil (in %)	Extended strand & beta turn (in %)	Helix (in %)	Random coil (in %)	Extended strand & beta turn (in %)
Chitinase1	12.07	62.54	25.39	22.60	57.89	19.50	24.15	70.90	4.95
Chitinase3	15.00	64.69	20.31	20.62	64.06	15.31	22.19	74.69	3.20
Chitinase5	15.97	59.72	24.31	27.78	51.74	20.49	30.21	67.71	2.08
Chitinase7	26.47	53.53	20.00	31.47	53.24	15.30	24.71	72.94	2.65
Chitinase8	18.39	58.62	22.99	33.33	44.06	22.61	35.25	64.75	0.00
Chitinase9	19.46	60.78	19.76	26.65	55.99	17.37	27.84	68.52	2.70
Chitinase10	21.72	58.28	20.00	27.24	49.31	23.45	31.03	64.83	4.14
Chitinase11	18.75	56.64	24.61	30.08	51.17	18.75	32.81	66.41	0.78
Chitinase12	19.63	62.58	17.79	25.15	59.20	15.64	27.00	71.47	1.53

Model building and refinement

The theoretical structure of rice chitinases were built using MODELLER-9v10 [48]. Modeller implements comparative protein structure modelling by satisfaction of spatial restraints. For 3-D model building of each chitinases the target sequences were aligned with their respective templates in Modeller. In this approach single-template and multiple-template based methods were used to build the 3-D structures. The models generated were subjected to loop refinement in Modeller.

Model evaluation

The qualities of the refined models were evaluated both geometrically and energetically by a series of tests for its internal consistency and reliability. We used PROCHECK [49], WHAT_CHECK [50], ERRAT [51] and VERIFY_3D [52, 53] tools embedded in the structure analysis and validation server (SAVES) (<http://nihserver.mbi.ucla.edu/SAVES/>). PROCHECK analyses the Ramachandran plot quality, peptide bond planarity, non bonded interactions, main chain hydrogen bond energy, C α chiralities and overall G factor. The non bonded interactions between different atom types were checked by ERRAT. Verify_3D was used to access the compatibility of the atomic model with its own amino acid sequence. A high Verify_3D profile score indicates the high quality of protein model. The Protein Structure Analysis (ProSA) (<https://prosa.services.came.sbg.ac.at/prosa.php>) [54] tool was employed in the refinement and validation of modelled structures. The root mean square deviation (RMSD) between the main chain atoms of models and respective templates were calculated by structural superimpositions of predicted structures with their respective templates using iPBA web server (http://www.dsimb.inserm.fr/dsimb_tools/ipba/index.php) [55]. To sum up the geometry, non-bonded interaction of atoms, energy profiles and RMSD of most of the models of rice chitinases are reasonable and reliable for further study.

Function prediction

The 3d2GO server (<http://www.sbg.bio.ic.ac.uk/phyre/pfd/index.html>) was used to predict functions of the validated models using sequence and structure in the reference of gene ontology (GO). Various sources of information like sequence homology to functionally annotated sequences, overall topological similarity to structures with known function and geometric and residue similarity of predicted functional sites to regions of known structures were used.

Identification of functional surface

For each modelled chitinases, we used the CASTp server (<http://sts.bioengr.uic.edu/castp/calculation.php>) [56] to

identify their functional surfaces, which is taken as the surface pocket containing annotated binding site residues. If there exist more than one such pocket on a protein structure, only the largest one is selected, as the largest pocket often corresponds to enzyme binding site [57].

Results and discussion

Twelve reviewed sequences of rice chitinases were retrieved from UniProtKB database belonging to class I, II and IV. Out of these sequences, chitinase1, 2, 3, 7, 9, 10, 12 belong to class I, chitinase8 and 11 belong to class II whereas chitinase4, 5 and 6 belong to class-IV sub-family. All of them are referred to as PR-3 protein, belonging to GH

Table 4 Residues involved in the disordered region of rice chitinases

Chitinase	Residues involved in disorder region predicted metaPrDOS server
Chitinase1	M1,R2,Q56,S57,Q58,C59,S60,G61,S62,C63,G64,G65,G66,G67,P68,T69,P70,P80,S81,G82,G83,G84,G85,S86,G87,V88,R319,P320,F321,N322,S323
Chitinase3	M1,R2,A3,Q54,S55,Q56,C57,S58,G59,G60,C61,G62,G63,G64,P65,T66,P67,P68,S69,S70,G71,G72,G73,S74,G75,V76,A77,S78,N313,Q314,R315,P316,Y317,P318,P319,S320
Chitinase5	M1,A2,N3,S4,P5,T6,P7,T8,R58,S59,G60,P61,C62,Y63,G64,G65,G66,G67,G68,G69,G70,G71,G72,G73,G74,G75,G76,G77,G78,G79,G80,G81,G82,S83,G84,V85,S86,V87,E88,G284,N285,L286,Y287,C288
Chitinase7	M1,I2,A3,A4,R5,A6,A7,N8,L9,Q68,S69,Q70,C71,R72,V73,S74,R75,D76,G77,G78,D79,D80,D81,V82,A83,A84,V85,L86,L87,T88,A89,P90,G91,G92,G93,R94,A95,G96,V97,A98,S99,D332,I333,M334,S335,A336,S337,A338,A339K340
Chitinase8	M1,T2,T3,T4,T5,T6,R7,F8,V9,Q10,L11,A12,A13,C14,A15,A16,A17,S18,L19,L20,A21,V22,A23,A24,S25,G26,A27,A28,A29,Q30,G31,V33,G33,N255,Q256,R257,N258,F259,A260,S261
Chitinase9	M1,K2,A3,T4,T5,T6,G57,C58,Q59,S60,Q61,C62,D63,G64,C65,G66,G67,G68,G69,G70,G71,G72,G73,G74,G75,G76,G77,G78,G79,G80,G81,G82,G83,G88,A85,V86,E87,A88,P325,F326,N327,S328,G329,L330,T331,A332,A333,Q334
Chitinase10	M1,A2,K3,P4,T5,P6,A7,P8,A31,R32,W33,Y34,G35,G36,G37,G38,G39,G40,G41,Y42,S43,P44,S45,P46,S47,P48,V49,S50,S51,E285,H286,Q287,Q288,P289,F290
Chitinase11	M1,R2,R3,L4,L5,I14,A15,A16,A17,G18,G19,A20,S21,G22,Q23,Q24,A25,G26,V27,G28,Q250,K251,P252,Y253,G254,G255,G256
Chitinase12	M1,R2,G55,C56,Q57,S58,Q59,C60,S61,A62,A63,G64,C65,G66,G67,G68,G69,P70,T71,P72,P73,S74,G75,S76,G77,G78,S79,G80,V81,A82,S83,G175,P176,D177,Q321,R322,P323,F324,G325,S326

Table 5 Domains and inter-linking linker of rice chitinase predicted using SBASE server

Chitinase	Type of domain	Domain name with sequence position from SBASE prediction	Linker length	% of G, S in the linker	Linker position
Chitinase1	Multiple domain	Chitin-binding, type 1 - like domain, 20–50 Glycoside hydrolase, family 19 - like domain, 86–321	35	34.28, 22.85	51–85
Chitinase3	Multiple domain	Chitin-binding, type 1 - like domain, 18–48 Glycoside hydrolase, family 19 - like domain, 84–309	35	32.35, 23.52	49–83
Chitinase5	Multiple domain	Chitin-binding, type 1 - like domain, 31–62 Glycoside hydrolase, family 19 - like domain, 92–288	29	74.07, 11.11	63–91
Chitinase7	Multiple domain	Chitin-binding, type 1 - like domain, 33–71 Glycoside hydrolase, family 19 - like domain, 100–323	28	21.42, 7.14	72–99
Chitinase8	Single domain	Glycoside hydrolase, family 19 - like domain, 36–253	-	-	-
Chitinase9	Multiple domain	Chitin-binding, type 1 - like domain, 24–53 Glycoside hydrolase, family 19 - like domain, 89–320	35	62.85, 2.85	54–88
Chitinase10	Single domain	Glycoside hydrolase, family 19 - like domain, 55–284	-	-	-
Chitinase11	Single domain	Glycoside hydrolase, family 19 - like domain, 30–247	-	-	-
Chitinase12	Multiple domain	Chitin-binding, type 1 - like domain, 22–52 Glycoside hydrolase, family 19 - like domain, 84–318	31	35.48, 19.35	53–83

family 19. Multiple sequence alignment was constructed in ClustalW which is shown in Fig. 1. The results showed that class I and class IV chitinases of rice have two common domains, i.e. N-terminal cysteine rich chitin binding domain (CBD) and a C-terminal catalytic domain (CatD). The aromatic amino acids belonging to CBD of class I and IV chitinases are highly conserved. Though chitinase10 belongs to class I, it lacks the N-terminus CBD which is a conserved feature of class I and class IV chitinases. CBD of class IV has a deletion of approximately seven amino acids residues in the N-terminal region. CatD possess several deletions of around 35 amino acids in the C-terminal region which may have resulted in the small molecular size of class IV chitinases in comparison to class I chitinases. Class II chitinases of rice are homologous to the CatD of class I and IV chitinases which lack the N-terminus CBD. Due to the lack of N-terminus CBD and few deletions in the CatD of class II chitinases, the

molecular size of class II chitinases is smaller than that of class I chitinases whereas there is a high level sequence similarity within the class. The two domains (CBD and CatD) of class I and class IV chitinases are interconnected by glycine, proline and serine rich linker peptide sequence.

A total of 74 sequences of chitinase from different plant species were selected and aligned in ClustalW for phylogeny analysis. The alignment was further used for phylogenetic tree construction using NJ method with a bootstrap value of 1000 in MEGA5.0. The consensus tree showed dichotomy with two different clusters (cluster I and II) having strong bootstrap value >70 with their nodes (Fig. 2a). On the basis of all chitinases analysed in the present study it may be surmised that the main clusters of chitinases in monocotyledons and dicotyledons differ from each other. Within the clusters of monocots their exist sub clusters (sub cluster Ia and Ib), indicating that the chitinases

Table 6 Model validation scores and energy of the models

Chitinase	Verify_3D	Errat	ProSA (Z score)	RMSD (Å°)	Energy of the model before energy minimisation (kcal/mol)	Energy of the model of after energy minimisation (kcal/mol)
Chitinase1	93.21	78.073	-6.36	0.48	-31518.003906	-32616.703125
Chitinase3	93.0	74.00	-6.83	0.59	-31305.187500	-32245.533203
Chitinase5	81.66	83.00	-5.39	0.43	-24716.453125	-25481.921875
Chitinase7	88.56	80.135	-5.52	0.59	-30552.615234	-31496.814453
Chitinase8	90.46	81.746	-5.32	0.60	-26796.945313	-27524.714844
Chitinase9	86.57	85.95	-5.28	0.75	-27245.490234	-28335.541016
Chitinase10	91.07	74.113	-5.65	0.97	-29054.037109	-30058.404297
Chitinase11	93.39	84.348	-6.57	0.68	-25638.656250	-26263.453125
Chitinase12	92.66	86.806	-6.47	0.77	-29809.345703	-30663.468750

in the same cluster are still evolving in different plants (Fig. 2b). The phylogeny analysis showed that despite sequence divergence among all chitinases of rice, they have evolved from a common ancestor and are conserved throughout the evolutionary process.

The physico-chemical properties of rice chitinases were computed using Protparam tool of Expsasy proteomic server is tabulated in Table 2. The molecular weight of the selected rice chitinases ranges between 27–35 KDa. Isoelectric point (pI) is the pH at which the surface of protein is covered with charge but net charge of protein is zero. At pI proteins are stable and compact. Computational results showed that

chitinase3, 8, 9, 11 and 12 are acidic in nature (pI <7.0) whereas chitinase1, 5, 7 and 10 are basic in nature (pI >7.0). The computed isoelectric point (pI) will be useful for developing buffer system for purification by isoelectric focusing method. The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chains is regarded as a positive factor for the increase of thermal stability of globular proteins [58]. Aliphatic index of chitinases ranges between 50.90–67.82. The very high aliphatic index of chitinases indicates that rice chitinase proteins may be stable for a wide range of temperature. The instability index provides an estimate of the stability of protein in a test tube.

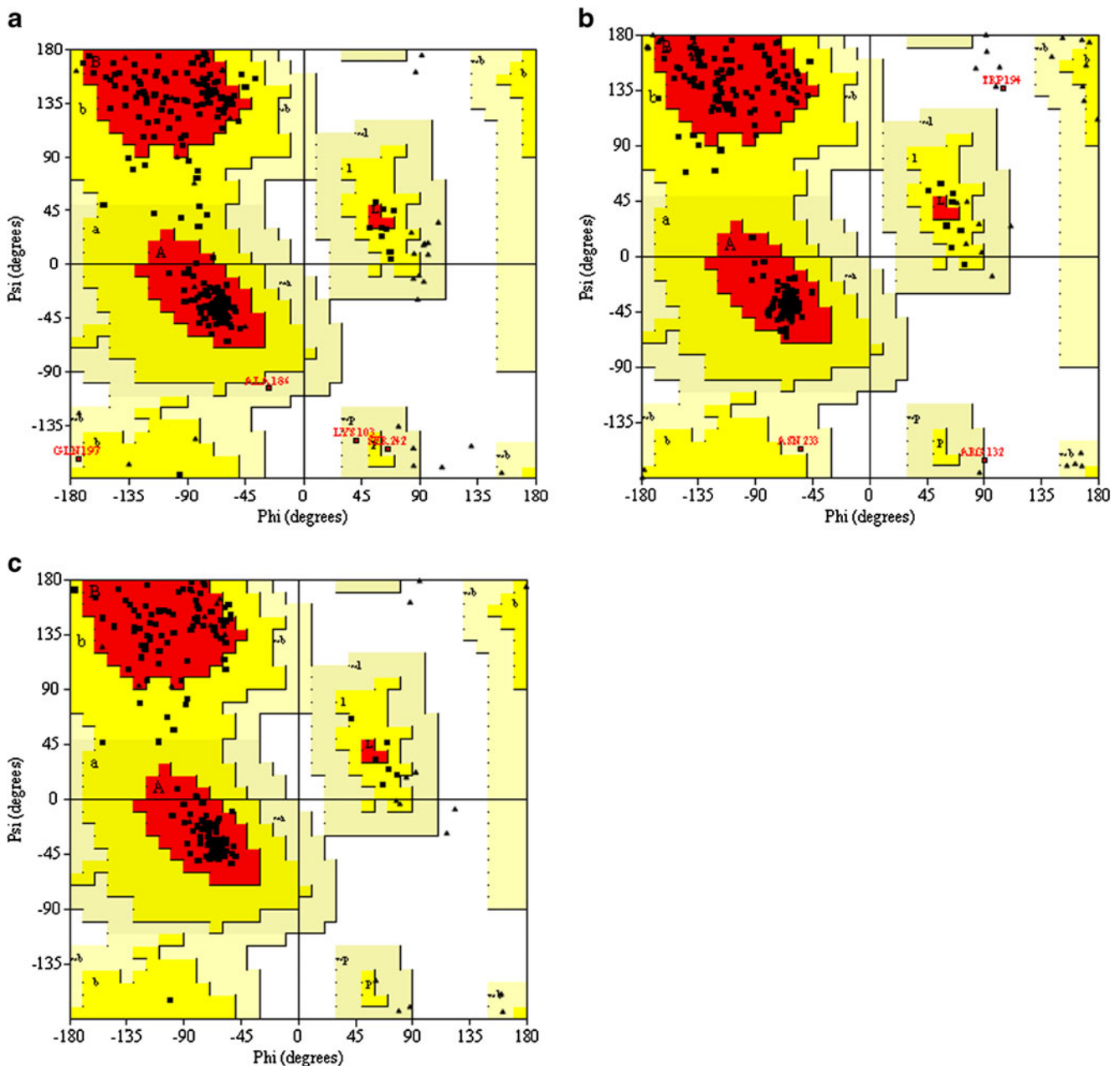


Fig. 3 Ramachandran plot of chitinases of rice. (a) Chitinase1 of class I (b) chitinase11 of class II (c) chitinase5 of class IV

There are certain dipeptides, the occurrence of which is significantly different in the unstable proteins compared with those in the stable ones. A protein whose instability index is smaller than 40 is predicted as stable, a value above 40 predicts that the protein may be unstable [59]. Except chitinase10, it was predicted that all chitinases are stable in nature showing an instability index <40. The grand average hydrophobicity (GRAVY) value for a peptide or protein is calculated as the sum of hydrophobicity values of all the amino acids, divided by the number of residues in the sequence [60]. GRAVY indices of chitinases range from -0.107 to -0.453 . This low range of value indicates the possibility of better interaction with water.

Secondary structures of chitinases were predicted using PSIPRED, SOPMA and GOR IV servers. The comparative study from all these servers showed that in all classes (I, II and IV) of chitinases investigated in the present study random coils dominated among secondary structure element followed by helices and strands which are shown in Table 3. Although these servers use different algorithms and approaches to predict secondary structure elements from primary amino acid sequences, results obtained from these servers are approximately the same and are considered for further investigation.

The metaPrDOS server was used to predict natively disordered regions of rice chitinases from its amino acid sequences. Proteins often in their native states have regions with very flexible and unstable structures, treated as disordered region which are involved in many biological processes such as regulation, signalling and cell cycle control [61, 62]. Disordered regions seem to be the molecular recognition site of proteins or DNA [63, 64]. During the interaction with the ligands, it is frequently observed that disordered regions transit to order where the flexibility of the region provides the high specificity and low affinity towards multiple partners [65]. Therefore, it is quite necessary to identify the disordered

regions of target proteins from their amino acid sequences. The results from this server revealed that few residues from the beginning of N-terminal regions and some residues at the end of C-terminal region fall in the disordered region. In addition, a stretch of consecutive residues in between the two termini of chitinases (the linker portion) fall in the disordered region. In all the classes (class I, II and IV) of rice chitinases the disordered regions are dominated by higher frequency of hydrophilic, charged residues, low sequence complexity regions and residues involved in phosphorylation (Serine, Threonine and Tyrosine). The residues predicted from the disordered region and their position are depicted in Table 4.

The domain boundaries of chitinases were predicted using SBASE server. It predicts domain boundaries as well as domain content and thus can be used for the identification of protein domain architecture. The results from this study showed that chitinase1, 3,5,7,9 and 12 are comprised of two domains namely chitin binding type-1 like domain and glycoside hydrolase, family 19-like domain whereas chitinase8, 10 and 11 possesses a single glycoside hydrolase, family 19-like domain. The typical domains of rice chitinases and their positions in the sequences are listed in Table 5. Discrete domains are often associated with multiple function of protein where domains are connected by inter-domain linkers. The linker region that is flexible in 3D-space, which may significantly affect the cooperation and interaction between domains, alter the overall functionality and efficiency of multiple domain proteins. They keep the domains apart and provide great extent of flexibility to move individually which is a part of their catalytic function. As chitinases possess discrete domains, it is important to predict the linker sequence which joins the adjacent domains present in it. The linker regions were manually curated and are reported in Table 5. We have analysed the amino acids propensities in linkers and examined their order of residues

Table 7 Ramachandran plot statistics and overall G- factors

Chitinase	Ramachandran plot (In %)				Overall G- factor
	Most favoured regions	Additional allowed regions	Generously allowed regions	Disallowed regions	
Chitinase1	88.8	9.6	1.6	0.0	-0.14
Chitinase3	88.5	9.9	1.6	0.0	-0.19
Chitinase5	90.2	8.5	0.9	0.4	-0.04
Chitinase7	89.2	9.0	1.4	0.4	-0.15
Chitinase8	86.7	11.9	1.4	0.0	-0.08
Chitinase9	90.3	9.3	0.4	0.0	-0.21
Chitinase10	85.9	11.5	2.6	0.0	-0.13
Chitinase11	90.7	9.3	0.0	0.0	-0.04
Chitinase12	88.9	10.0	1.1	0.0	-0.18

within linkers. The amino acid glycine (G), proline (P) and serine (S) are more predominant in these linker regions which may provide flexibility to the two discrete domains present in rice chitinase to act independently. It was also reported that penta-peptides consisting of Gly, Ser and Thr would make the best linkers for gene fusion; as these residues were most strongly preferred within natural linkers [66].

Comparative modelling of protein is considered as one of the most accurate methods for 3D structure- function prediction, yielding suitable models for wide spectrum of applications [67]. It is usually a method of choice when a clear relationship of homology between the sequences of target protein and at least

one known structure is found. The approach would give reasonable results based on the assumption that the tertiary structure of two proteins will be similar if their sequences are related [68]. Rice possesses several family 19 chitinases but only class I chitinase (*OsChia1b*), referred to as Cht-2, has been reported in PDB (PDBID: 2DKV) to date. Therefore it prompted us to construct the homology models of different classes of Family 19 chitinases of rice, the structures of which have not been reported. Templates were retrieved by performing BLAST search against PDB. Templates were selected based on the query coverage, sequence identity, less E-value and structural resolution. For comparative modelling the single-template and

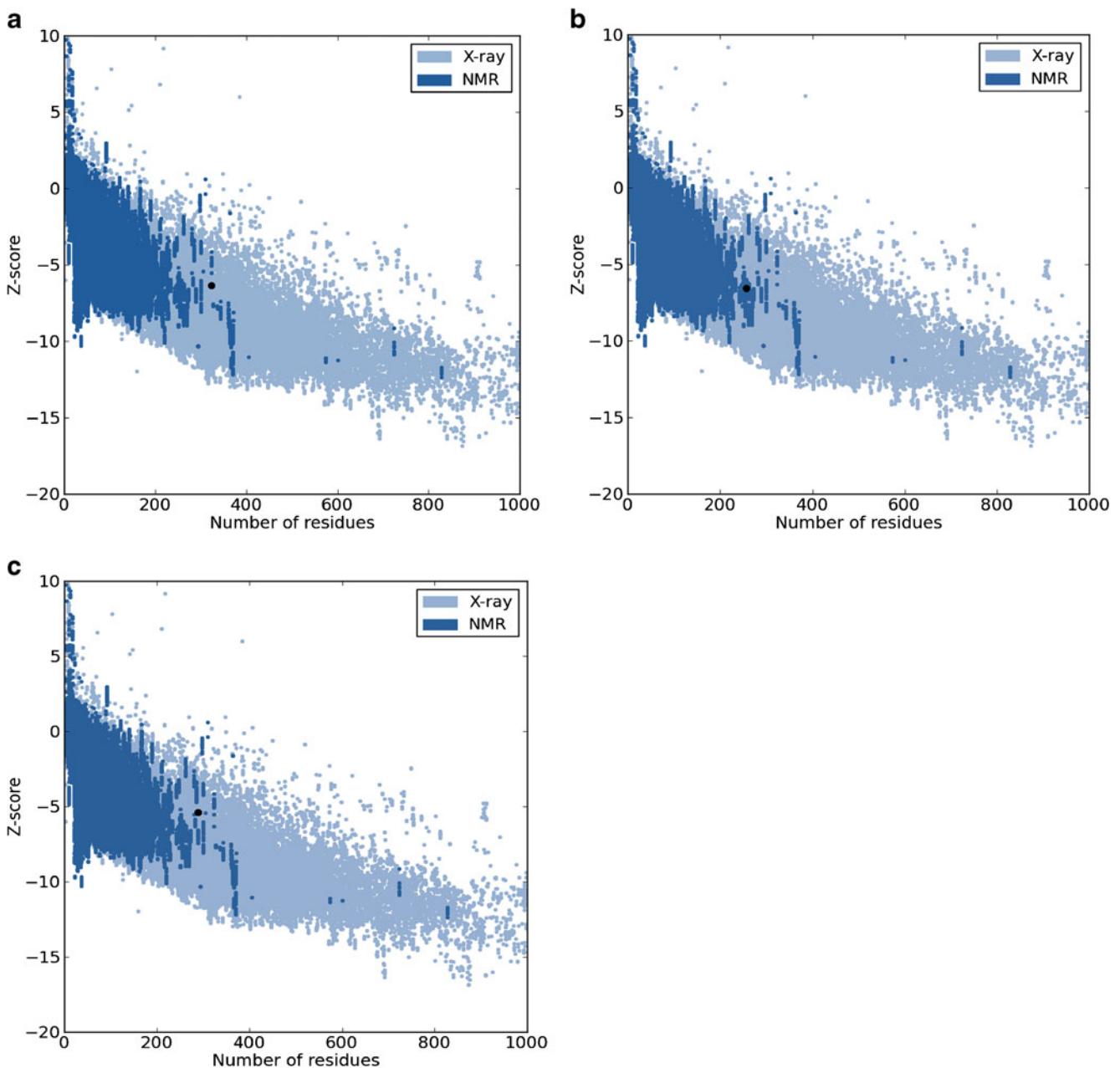


Fig. 4 ProSA energy profile of rice chitinase. (a) Chitinase1 of class I (b) chitinase11 of class II (c) chitinase5 of class IV

multi-template approaches were used. 2DKV and 2Z39 were considered the best templates for model building of chitinase1, 2, 9 and 12 as they showed high percentage of identity with the query sequences, 2UVO and 3HBD was chosen as template for chitinase5. 2BAA was used as template for chitinase11. For chitinase8, and 10 multiple templates were selected as it is recommended to use multiple templates (when available) to

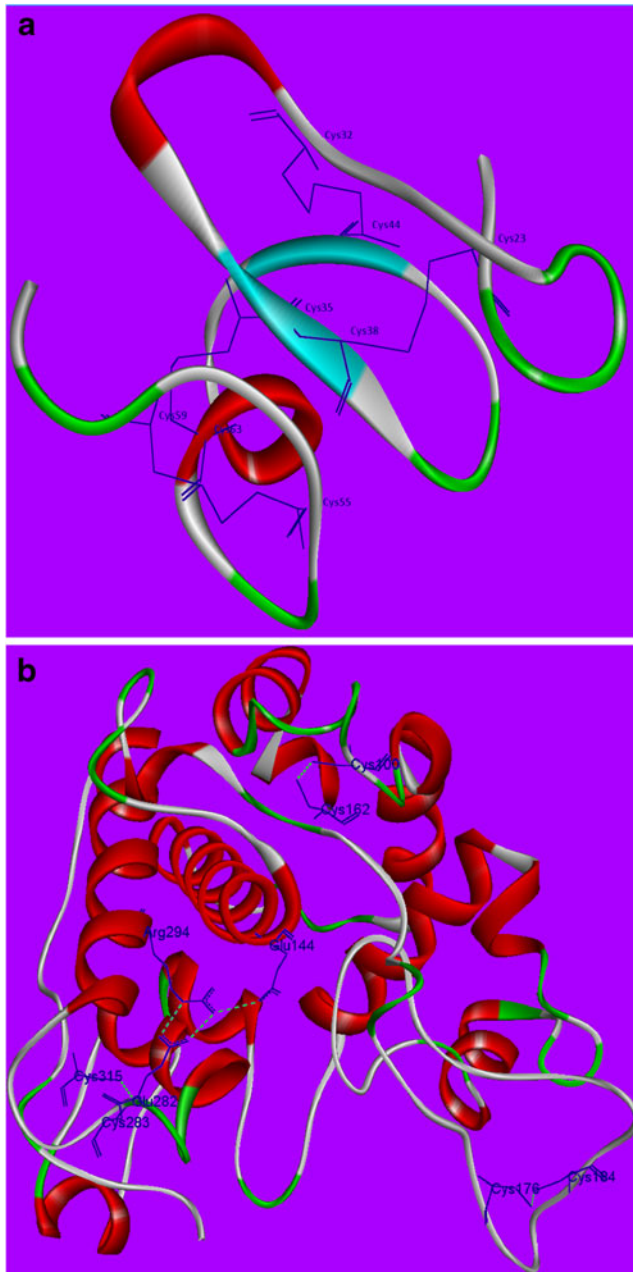


Fig. 5 Homology model of chitinase1 from rice. (a) Solid ribbon representation of the chitin binding domain (CBD) coloured by their secondary structure elements and disulfide bridges between cysteine residues are labelled (b) Solid ribbon representation of the catalytic domain (CatD) coloured by their secondary structure elements. Disulfide bridges between cysteine residues and conserved residues are labelled

avoid biasing the models towards one protein or one set of side chain conformations [69, 70].

Modeller9v10 was used to build the three-dimensional models of chitinases based on the target-template alignment. Modeller generated five predicted structures for each rice chitinase. The models with the lowest discrete optimized protein energy (DOPE), (a statistical potential used to assess homology models) scores were considered to be thermodynamically stable and chosen for further refinement and validation. The models with amino acids in the loop region were subjected to loop refinement in Modeller using loop_refine.py script. The loop refinement script generated five different models of chitinase, the models with lowest energy were chosen for further study. The total energy of models was calculated by GROMOS96 force field and energies of the models before and after refinement in Modeller are depicted in Table 6. Decrease in force field energies after refinement confirms that models were refined. Minimum energy values possessed by almost all models compared to templates indicates stability of all models.

The overall stereo-chemical quality and accuracy of the predicted models was evaluated using Ramachandran plot (Fig. 3) in Procheck. The refined models showed good percent of residues in most favoured regions, additional allowed regions and generously allowed regions (depicted in Table 7). Absence of residues from disallowed regions in chitinase (except chitinase5 and chitinase7) supports its high geometric quality. Though chitinase 5 and chitinase 7 models have one residue falling in disallowed region of Ramachandran plot,

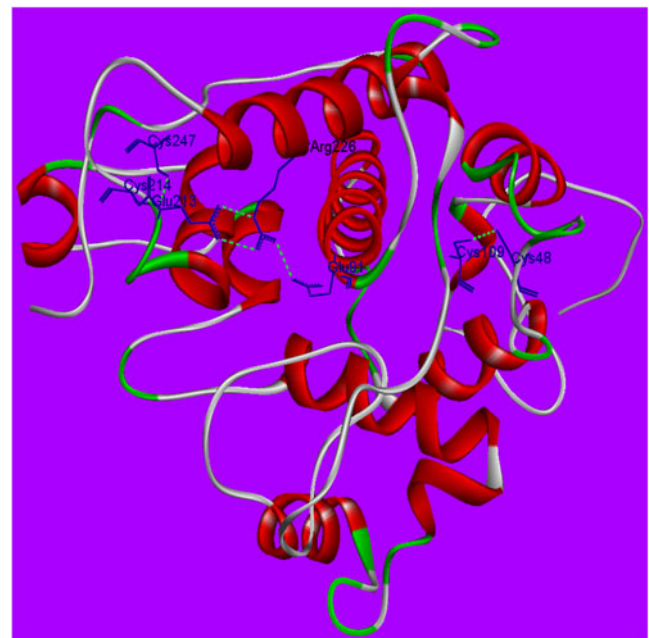


Fig. 6 Solid ribbon representation of the catalytic domain (CatD) of chitinase11 coloured by its secondary structure elements. Disulfide bridges between cysteine residues and conserved residues are labelled

they do not interfere with active sites, hence these models are acceptable. The total quality (G factors) was also obtained in acceptable range as shown in Table 7 (acceptable values of G-factor in Procheck are between 0 and -0.5 with the best model displaying values close to zero) [60] indicated the designed models are of good quality and acceptable.

The packing quality of each residue as assessed by the Verify_3D program represents the profile obtained with respect to the residues. Compatibility of the model residues with their environment is assessed by a score function. Residues with a score over 0.2 should be considered reliable. Score for all refined models maximally lies above 0.2 which corresponds to acceptable side chain environment as represented in Table 6.

Energy profiles of models were obtained using ProSA score (Fig. 4). ProSA revealed a Z-score (a measure of quality of model as it measures the total energy of the structures) value of the model which lies between -5.0 to -7.5 (negative value imply model accuracy) as depicted in Table 6. The degree of structure similarity was measured using root-mean-square distance (RMSD) between equivalent atom pairs. To investigate how well the modelled structure matches the X-ray data of template, the prepared models and their respective templates were superimposed on their backbone atoms. RMSD values of the backbone atoms for all models tabulated in Table 6 supported that generated models are reasonably good and quite similar to template. RMSD analysis of the chitinase models was measured from its templates using iPBA web server. The coordinates of the models are deposited in Protein Model DataBase (PMDb) and can be accessed at <http://mi.caspu.it/PMDb> using PMDB ID: PM0077946-954. Refined and validated models are shown in Figs. 5, 6 and 7 respectively.

The structures of each chitinases belonging to three different classes were analysed extensively. The predicted structure of class I chitinases (chitinase1, 3, 7, 9 and 12) of rice revealed that they are composed of two discrete domains which are interlinked by hinge region rich in proline and glycine with an exception of chitinase10, which do not possess the N-terminal CBD. This correlates with the comparative sequence analysis. The secondary structure elements predicted by different servers correlate significantly with the results from STRIDE [71], where stride recognises secondary structural elements in protein from their atomic coordinates. The two discrete domains (CBD and CatD) were built independently and were analysed.

The 3-D structure of CBD of chitinase1 was modelled by considering the CBD of its nearest neighbour 2DKV (chitinase2) as template. The predicted structure of chitinase1 belonging to class I chitinase revealed that the CBD is composed of a 3_{10} helix, an α -helix and a two stranded anti-parallel β -sheets ((Cys37-Ser39) and (Asn43-Gly45)) which are connected by a turn composed of three residues (Gln40-Gly42). The CBD is connected to CBD by flexible linker sequence rich in glycine and proline residues. A total of three disulfide bridges (Cys23/Cys38, Cys32/Cys44 and

Cys35/Cys51) maintain the shape and stability of the fold which are conserved in related sequences of class I family. The aromatic residues tyrosine and tryptophan are highly conserved which are mostly clustered on one face of the protein. The CBD also possess cysteine and glycine residues at conserved positions. The conserved cysteine residues forming disulfide bridges are thought to maintain the stability of the structure which may help the enzyme in extracellular activities and conserved glycine residues may help correct folding of the chitinase.

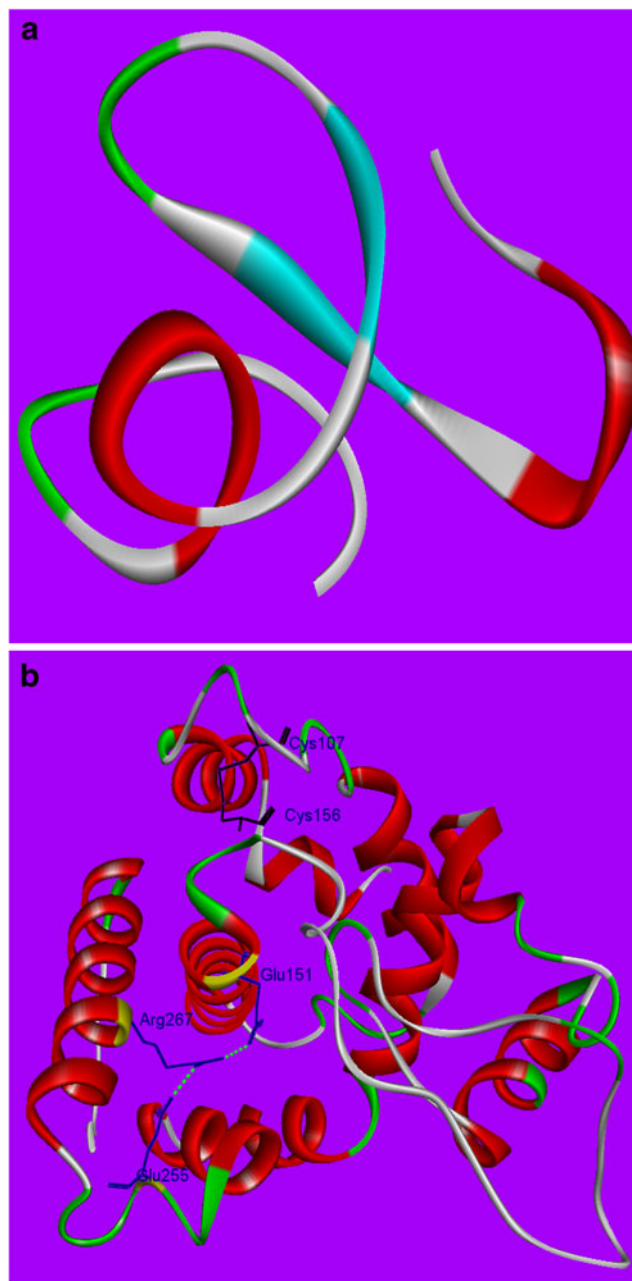


Fig. 7 Solid ribbon representation of the CBD (a) and CatD (b) of chitinase5 coloured by its secondary structure elements. Disulfide bridges between cysteine residues and conserved residues of CatD are labelled

The CatD of chitinase1 was modelled by considering CatD of mustard chitinase (2Z39) as template. The CatD of chitinase1 of rice is dominated by α -helix, comprised of 12 helices and five disulfide bridges. The structure resembles that of CatD of mustard and barley chitinases (GH family 19). The superposition (overlapping of C α atoms) of CatD with that of corresponding domain of 2Z39 reveals the spatial position and orientation of the helices are highly conserved (Fig. 8b). CatD of chitinase1 has a triad consisting of Glu144, Glu282 along with Arg294 which corresponds to the catalytic triad of its template 2Z39 (Glu212,

Glu349 along with Arg361). Ubhayasekera et al. [72, 73] have reported Arg361 and Glu349 residues are essential for catalysis and works together along with Glu212 in the form of a catalytic triad in 2Z39. As evidenced from the structural alignment between chitinase1 with 2Z39, the key elements of secondary structure are strongly conserved (secondary structure elements identity of 93.8 %). From Fig. 8b, it is evident that the structure of CatD possesses five loops namely I, II, III, IV and V, which are the most important striking features of class-I and II family 19 chitinases. Loops I, II and V are missing in bacterial chitinases [74]. The

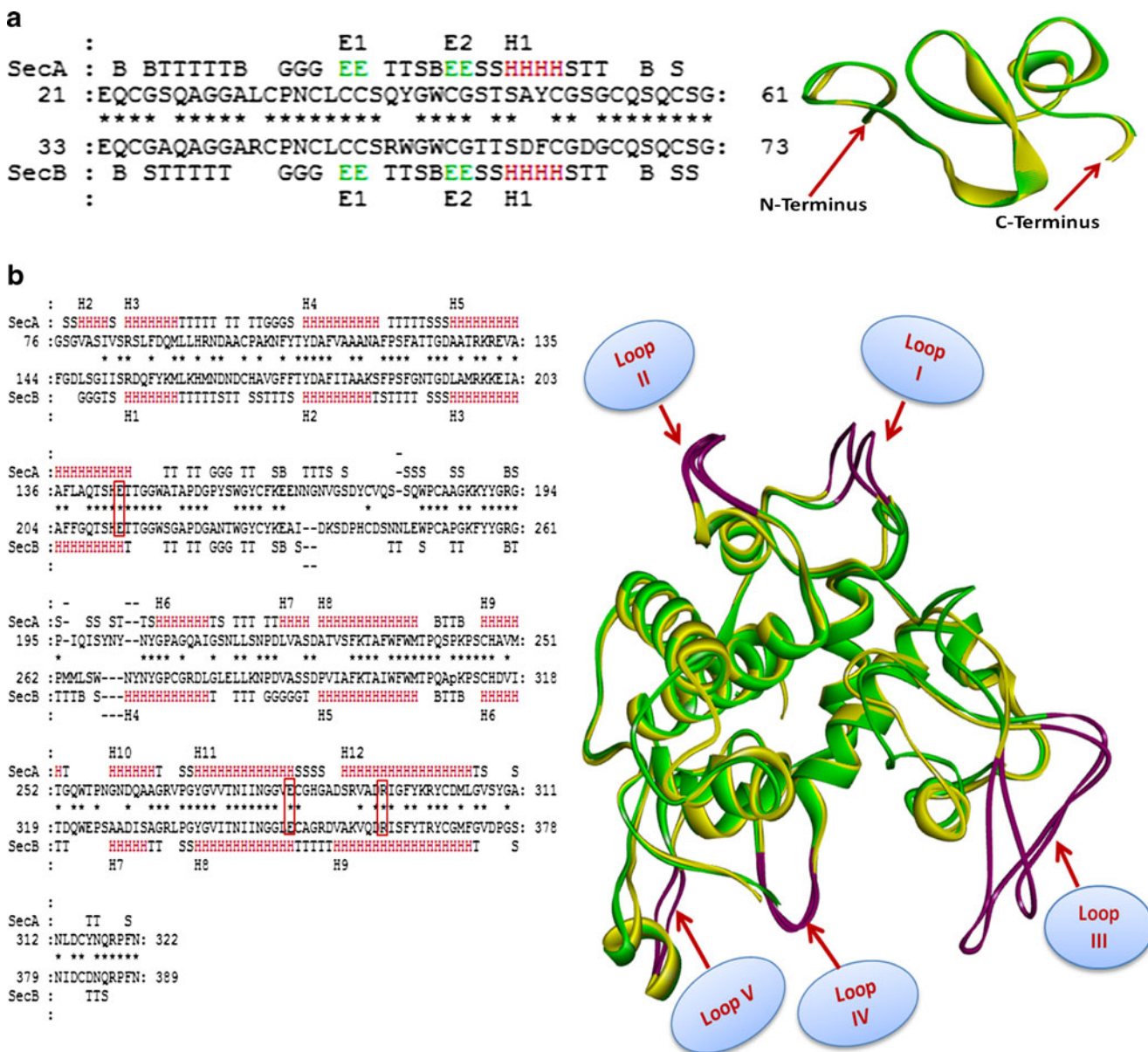


Fig. 8 Structural superposition of chitinase1 with its template. (a) Superposition of CBD with 2DKV along with their secondary structure alignment [* Sec A: Chitinase1, Sec B: 2DKV]. CBD of chitinase1 is coloured in green and CBD of template is coloured in yellow. (b) Superposition of CatD with 2Z39 along with their secondary structure

alignment [* Sec A: Chitinase1 Sec B: 2Z39]. CatD of chitinase1 is coloured in green and CatD of template is coloured in yellow. The loops are coloured in red. The key conserved residues are highlighted in red square box

missing loop information in family 19 reflects that loop III which is common in all classes of chitinase might play an important role and probably has a significant function. The structure of chitinase 3 is exactly the same as chitinase1 whereas structures of chitinase7, 9, 10 and 12 almost resemble that of chitinase1 and 3. There is a small difference in their number of α -helices and the disulfide bridges which help in maintaining the stability of the folds present in chitinase of rice.

Chitinase11 belonging to class II of GH family 19 reveals that the structure possess only the CatD and lacks the CBD. The CatD of chitinase11 is almost the same as that of class I chitinase and is dominated by α helices. The two disulfide bridges (Cys48/Cys109 and Cys214/Cys247) maintain overall stability of the bilobed structure. The loops (I, II, III, IV and V), which are also common in class I chitinases are located in CatD of chitinase11. As evidenced from the structural alignment between chitinase11 with its template 2BAA (Figs. 9 and 10) the key elements of secondary structure are strongly conserved (secondary structure elements identity of 96.9 %). The key catalytic residues of template (2BAA), Glu67 super imposes with Glu91,

Glu89 with Glu113 and (Tyr123/Tyr133, Asn124/Asn134, Gln118/Gln128, Gln162/Gln172, Lys165/Lys175, Pro163/Pro173 and Asn199/Asn209) of chitinase11, and are fully conserved. The structure of chitinase8 is also the same as that of chitinase11, the only difference being observed in their number of alpha helices (12 in the case of chitinase11 and 13 in the case of chitinase8).

Chitinase5 belonging to class IV chitinase, revealed that it is somewhat different from class I and II chitinases and dominated by α -helices. It is also seen that it has fewer disulfide bridges within the structure in comparison to class I and II chitinases. The two domains (CBD and CatD) possessed by chitinase5 were modelled independently with their closest neighbouring structure as template. The CBD of chitinase5 was built by homology modelling using the closest homologous structure, wheat germ agglutinin complex with n-acetyl-d-glucosamine (PDB ID: 2UVO B chain) as template. The CBD of this chitinase is globular in shape, very small in size and has irregular β -sheets with two strands. It is connected to the catalytic module by a flexible linker. The surface aromatic residues which are concentrated on the surface of the protein are highly conserved.

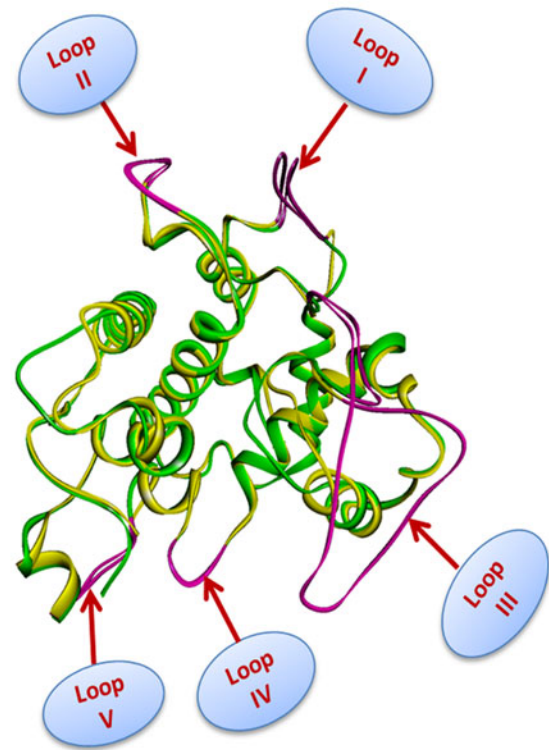
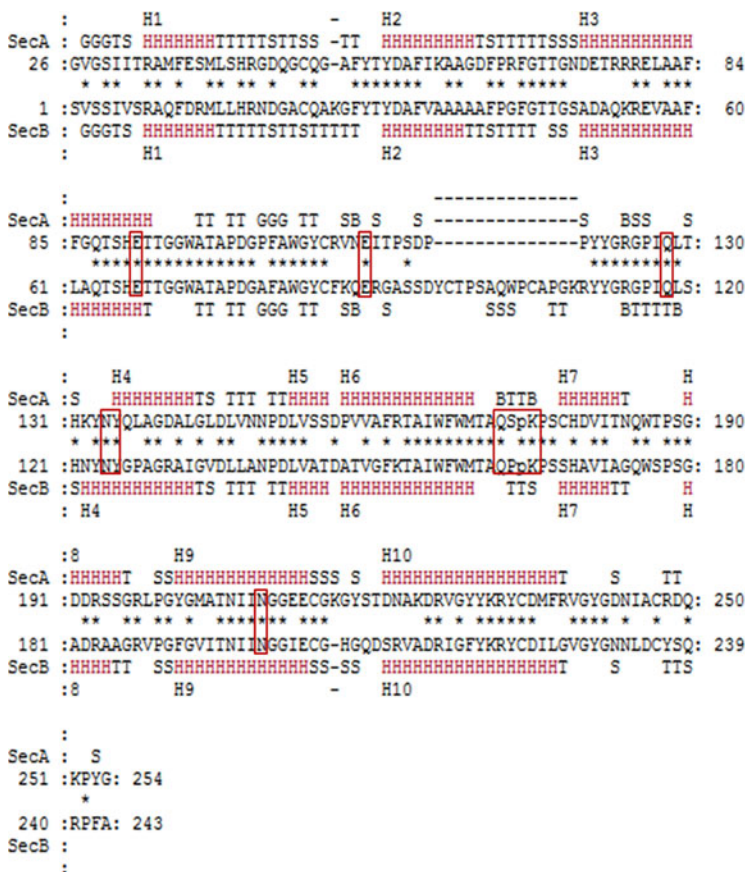


Fig. 9 Structural superposition of chitinase11 with its template 2BAA along with their secondary structure alignment [* Sec A: Chitinase11 Sec B: 2BAA]. CBD of chitinase11 is coloured in green and CBD of template

is coloured in yellow. The loops are coloured in red. The key conserved residues are highlighted in red square box

Table 8 The active sites and the amino acids involved in the formation of the large cavity in chitinases with their surface areas and volumes predicted by CastP server

Chitinase	Pocket information		Amino acids involved in the cavity
	Area (Sq. A°)	Volume (Cubic A°)	
Chitinase1	1902.4	2427.1	Y106,F137,Q140,T141,E144,T145,Y161,F163,K164,E165,E166,Y175,C176,V177,S179,W182,C184,K189,Y190,Y191,G192,R193,G194,P195,I196,Q197,I198,S199,Y200,N201,Y202,N203,Y204,L215,L216,P219,D220,V222,A223,S229,F230,T232,A233,W235,F236,W237,T239,Q241,S242,K244,P245,S246,C247,N275,I276,I277,N278,G278,G280,R294,N322
Chitinase3	639.2	2899.3	M1,R2,A3,L4,A5,A7,V8,V9,A10,A12,V13,V14,A15,V16,R17,G18,E19,Q20,Q24,A25,A28,L29,P31,S81,S83,L84,Q87,D299,M300,L301
Chitinase5	214.8	276.3	E151,L190,Q91,I192,S193,W194,N197,F230,I250,N251
Chitinase7	1420.9	2951.9	A19,V20,L21,L23,A24,Y25,A26,A27,A28,T29,A30,R31,A32,E33,P45,N46,R47,L48,C49,C56,G57,L58,T59,C63,Q70,C71,R72,V73,D80,D81,V82,A83,A84,V85,L87,T88,A89,P90,G91,G92,G93,R94,A95,S103,D104,E107,L110,P111,H112,R113,D114,D115,G122,Y124,A125,Y126,R127,V130,A131
Chitinase8	898.9	1470.4	T3,T6,R7,F8,V9,Q10,L11,A12,A13,C14,A16,A17,S18,A,L20,A21,V22,A23,A24,S25,G26,A27,A28,Q29,G30,G31,V32,G33,S34,V35,T37,Q38,A39,I65,A68,N69,F71,P72,A73,F74,G75,T76,G78,G79,S80,A81,I84,R85
Chitinase9	2307.0	5189.0	M1,A3,T5,T6,A7,V8,A9,L11,A13,A14,A15,A16,M17,G27,G32,A33,L34,C41,C58,Q59,S60,Q61,C62,D63,G64,C65,G66,G67,G68,G69,G70,G71,G72,G82,G83,G84,A85,V86,E87,A88,V89,V90,F95,E96,Q97,L98,L99,L100,H101,R102,N103,D104,A105,C107,A109,R110,G111,F112,Y113,T114,Y115,D116,L118,V119,T120,A122,A123,F128,A129,A130,R137,K138,V141,F144,L145,S149,P160,D161,S165,W166,G167,Y168,C169,F235,L239,R305,Y306,V309,L310
Chitinase10	518.5	550.1	G161,R162,G163,I165,Q166,N170,F171,N172,Y173,G174,A176,G177,L180,F182,G184,N187,P188,I190,V191,I197,T201,A202,W204,F205
Chitinase11	396.6	481.9	E113,Y121,G123,R124,G125,P126,I127,Q128,L129,T130,H131,K132,Y133,N134,Y135,A138,L146,V153,T163,A164,W166,F167
Chitinase12	1772.1	3912.3	M9,V10,A11,T12,A13,F14,L15,A16,A18,G30,V32,C33,P34,N35,C36,L37,C38,S47,C52,S58,Q59,C60,S61,A62,A63,G64,C65,G67,G68,G69,P70,T71,P72,P73,S74,S76,G77,G78,S79,G80,V81,A82,S83,I84,V85,S86,R87,S88,L89,N106,F107,T109,Y110,D111,A112,V114,A115,A117,S118,F120,P121,G122,F123,A124,A125,A126,G127,G128,D129,A130,D131,N133,K134,R135,E136,V137,S227,D228,A229,T230,H251,A254,T255,G256,Q257,L308,V310

predominantly associated with different cellular process, i.e. cell wall macromolecules catabolic process, chitin catabolic process, chitinase activity, hydrolase activity, hydrolysing O-glycosyl compounds, hydrolase activity acting on glycosyl bond with a confidence value >0.85 whereas model of chitinase12 showed GO terms associated with polysaccharide binding, pattern binding and carbohydrate binding with a lower confidence value (<0.25). The multiplicity of function in different classes of chitinase predicted by 3d2GO server is perfectly justified by structural variation in all forms of chitinases in rice.

Structural pockets and cavities are often associated with the binding sites and active sites of proteins respectively [56]. The active sites and the amino acids involved in the formation of the large cavity with their surface areas and volumes are depicted in Table 8. Most of the residues that form the large cavities are hydrophilic and charged in nature, which are also conserved in all forms of chitinase in rice. The study of active sites reveals that the catalytic domain possesses few aromatic amino acids. Therefore CatD is dominated by hydrogen bonding in contrast to other interactions.

Conclusions

Chitinases are prime molecules of interest of plant pathologists and can be used in a variety of ways to improve plant health. These enzymes are not only involved in plant resistance to external environmental factors by generating signal molecules but also in plant growth and development. They are classified into various types on the basis of the structural and functional properties. The structure and function study gives a brief outline about enzyme 3-D structure and displays how the secondary structure elements are arranged in the protein. Rice chitinases are single or multi domain and multi functional enzymes. The evolutionary analysis revealed how the enzyme evolved during the process of evolution and its relatedness among other plant chitinases. The structural similarity as well as their differences among three different classes of chitinases (I, II and IV) from rice is reported in this work. The different class of chitinases belonging to GH family 19 possess highly α -helical and bilobed structures in nature. The superposition of all the classes of chitinases along with their closest homologous templates

reflects the secondary structure elements are strongly conserved. The highly conserved residues are catalytic in nature, help in substrate binding as well in disulfide bridge formation. One of the most important striking features of the three different classes of chitinase of rice is that the CatD possesses a catalytic triad which is thought to be involved in catalytic process common in all forms. As far as the loops formed within the CatD are concerned, class I and II chitinases of rice possess loop I, II, III, IV and V whereas loop II, IV and V are missing in class IV chitinase. Loop III which is common in all classes of chitinases might play an important role in their respective function. Our study also confirms that the absence and presence of different loops in GH family 19 of rice may be responsible for various sized products as previously reported by Mizuno et al. and Fukamizo et al. [75, 76]. So it can be concluded, that the sequence variation in different forms of chitinase might lead to the structural variation which reflects in terms of multiple functions which also fits with the prediction made by 3d2GO server for function prediction. The study of active sites reveals that the CatD is dominated by hydrogen bonding as only a few aromatic amino acids lie in the active sites for interaction. More structural study of this enzyme from different plants may enhance the knowledge of catalytic mechanism and substrate binding.

Acknowledgments The authors thankfully acknowledge the financial support for Agri-Bioinformatics Promotion Program by Bioinformatics Initiative Division, Department of Information Technology, Ministry of Communications & Information Technology, Government of India, New Delhi as well as Biotechnology Information System Network, Department of Biotechnology, Government of India. Also the authors are grateful to Assam Agricultural University, Jorhat, Assam for providing the necessary facilities, constant support to carryout our research endeavour.

References

- Bradley D, Kjellbom P, Lamb C (1992) Elicitor- and wound-induced oxidative cross-linking of a proline-rich plant cell wall protein: a novel, rapid defense response. *Cell* 70:21–30
- Hammerschmidt R (1999) Phytoalexins: what have we learned after 60 years? *Annu Rev Phytopathol* 37:285–306
- Misra AK, Gupta VK (2009) Trichoderma: biology, biodiversity and biotechnology. *J Eco-Friendly Agric* 4:99–117
- Hahlbrock K, Lamb CJ, Purwin C, Ebel J, Fautz E, Schafer E (1981) Rapid response of suspension-cultured parsley cells to the elicitor from *Phytophthora megasperma* var. *sojae*: induction of the enzymes of general phenylpropanoid metabolism. *Plant Physiol* 67:768–773
- Cramer CL, Ryder TB, Bell JN, Lamb CJ (1985) Rapid switching of plant gene expression induced by fungal elicitor. *Science* 227:1240–1243
- Lin W, Anuratha CS, Datta K, Potrykus I, Muthukrishnan S, Datta SK (1995) Genetic engineering of rice for resistance to sheath blight. *Nat Biotechnol* 13:686–691
- Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U, Vad K (1993) Plant chitinases. *Plant J* 3:31–40
- Regalado AP, Pinheiro C, Vidal S, Chaves I, Ricardo CP, Rodrigues-Pousada C (2000) The *Lupinus albus* class-III chitinase gene, IF3, is constitutively expressed in vegetative organs and developing seeds. *Planta* 210:543–550
- Hamel F, Bellemare G (1995) Characterization of a class I chitinase gene and of wound-inducible, root and flower-specific chitinase expression in *Brassica napus*. *Biochim Biophys Acta* 1263:212–220
- Krishnaveni S, Liang GH, Muthukrishnan S, Manickam A (1999) Purification and partial characterization of chitinases from sorghum seeds. *Plant Sci* 144:1–7
- Goormachtig S, Lievens S, Van de Velde W, Van Montagu M, Holsters M (1998) Srch13, a novel early nodulin from *Sesbania rostrata*, is related to acidic class III chitinases. *Plant Cell* 10:905–915
- Cullimore JV, Ranjeva R, Bono JJ (2001) Perception of lipochitooligosaccharidic Nod factors in legume. *Trends Plant Sci* 6:24–30
- van der Holst PP, Schlaman HR, Spaink HP (2001) Proteins involved in the production and perception of oligosaccharides in relation to plant and animal development. *Curr Opin Struct Biol* 11:608–616
- van Hengel AJ, Guzzo F, van Kammen A, de Vries SC (1998) Expression pattern of the carrot EP3 endochitinase genes in suspension cultures and in developing seeds. *Plant Physiol* 117:43–53
- Passarinho PA, Van Hengel AJ, Fransz PF, de Vries SC (2001) Expression pattern of the Arabidopsis thaliana AtEP3/AtchitIV endochitinase gene. *Planta* 212:556–567
- Cohen-Kupiec R, Chet I (1998) The molecular biology of chitin digestion. *Curr Opin Biotechnol* 9:270–277
- Hanfrey C, Fife M, Buchanan-Wollaston V (1996) Leaf senescence in *Brassica napus*: expression of genes encoding pathogenesis-related proteins. *Plant Mol Biol* 30:595–609
- Bishop JG, Dean AM, Mitchell-Olds T (2000) Rapid evolution in plant chitinases: molecular targets of selection in plant-pathogen coevolution. *Proc Natl Acad Sci U S A* 97:5322–5327
- Kasprzewska A (2003) Plant chitinases-regulation and function. *Cell Mol Biol Lett* 8:809–824
- Neuhaus JM, Fritig B, Linthorst HJM, Meins F Jr, Mikkelsen JD, Ryals J (1996) A revised nomenclature for chitinase genes. *Plant Mol Biol Rep* 14:102–104
- Arie M, Hikichi K, Takahashi K, Esaka M (2000) Characterisation of basic chitinase which is secreted by cultured pumpkin cells. *Physiol Plant* 110:232–239
- Melchers LS, Apotheker-de Groot M, van der Knaap JA, Ponstein AS, Sela-Buurlage MB, Bol JF, Cornelissen BJ, van den Elzen PJ, Linthorst HJ (1994) A new class of tobacco chitinases homologous to bacterial exo-chitinases displays antifungal activity. *Plant J* 5:469–480
- Jekel PA, Hartmann BH, Beintema JJ (1991) The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. *Eur J Biochem* 200:123–130
- Lerner DR, Raikhel NV (1992) The gene for stinging nettle lectin (urtica dioica agglutinin) encodes both a lectin and a chitinase. *J Biol Chem* 267:11085–11091
- Zhao KJ, Chye ML (1999) Methyl jasmonate induces expression of a novel Brassica juncea chitinase with two chitin-binding domains. *Plant Mol Biol* 40:1009–1018
- Berglund L, Brunstedt J, Nielsen KK, Chen Z, Mikkelsen JD, Marcker KA (1995) A proline-rich chitinase from *Beta vulgaris*. *Plant Mol Biol* 27:211–216
- Truong NH, Park SM, Nishizawa Y, Watanabe T, Sasaki T, Itoh Y (2003) Structure, heterologous expression, and properties of rice (*Oryza sativa* L.) family, 19 chitinases. *Biosci Biotech Biochem* 67:1063–1070

28. Kezuka Y, Kojima M, Mizuno R, Suzuki K, Watanabe T, Nonaka T (2010) Structure of full-length class I chitinase from rice revealed by X-ray crystallography and small-angle X-ray scattering. *Proteins* 78:2295–2305
29. Ouyang SW, Zhao KJ, Feng LX (2001) The structure and function, classification and evolution of plant chitinases. *CBB* 18:418–426
30. Aronson NN Jr, Halloran BA, Alexyev MF, Amable L, Madura JD, Pasupulati L, Worth C, Van Roey P (2003) Family 18 chitinase-oligosaccharide substrate interaction: subsite preference and anomer selectivity of *Serratia marcescens* chitinase A. *Biochem J* 376:87–95
31. Ohno T, Armand S, Hata T, Nikaidou N, Henrissat B, Mitsutomi M, Watanabe T (1996) A modular family chitinase found in the prokaryotic organism *Streptomyces griseus* HUT6037. *J Bacteriol* 178:5065–5070
32. Xu F, Fan C, He Y (2007) Chitinases in *Oryza sativa* ssp. *japonica* and *Arabidopsis thaliana*. *J Genet Genomics* 34:138–150
33. Nishizawa Y, Nishio Z, Nakazono K, Soma M, Nakajima E, Ugaki M, Hibi T (1999) Enhanced resistance to blast (*Magnaporthe grisea*) in transgenic Japonica rice by constitutive expression of rice chitinase. *Theor Appl Genet* 99:383–390
34. Datta K, Tu J, Oliva N, Ona I, Velazhahan R, Mew TW, Muthukrishnan S, Datta SK (2001) Enhanced resistance to sheath blight by constitutive expression of infection-related rice chitinase in transgenic elite indica rice cultivars. *Plant Sci* 160:405–414
35. Park SM, Kim DH, Truong NH, Itoh Y (2002) Heterologous expression and characterization of class III chitinases from rice (*Oryza sativa* L.). *Enzyme Microb Technol* 30:697–702
36. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
37. Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406–425
38. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
39. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy Server. In: John M. Walker (ed) *The proteomics protocols handbook*. Humana Press pp 571–607
40. Buchan DW, Ward SM, Loble AE, Nugent TC, Bryson K, Jones DT (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Res* 38:W563–W568
41. Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* 11:681–684
42. Garnier J, Gibrat JF, Robson B (1996) GOR secondary structure prediction method version IV. *Methods Enzymol* 266:540–553
43. Ishida T, Kinoshita K (2008) Prediction of disordered regions in proteins based on the meta approach. *Bioinformatics* 24:1344–1348
44. Kong L, Ranganathan S (2004) Delineation of modular proteins: domain boundary prediction from sequence information. *Brief Bioinform* 5:179–192
45. Vlahovicek K, Kaján L, Agoston V, Pongor S (2005) The SBASE domain sequence resource, release 12: prediction of protein domain-architecture using support vector machines. *Nucleic Acids Res* 33:D223–D225
46. Wootton JC (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem* 18:269–285
47. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
48. Sali A, Potterton L, Yuan F, van Vlijmen H, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. *Proteins* 23:318–326
49. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK—a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291
50. Hoof RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* 381:272
51. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
52. Bowie JU, Luthy R, Eisenberg D (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253:164–170
53. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356:83–85
54. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
55. Gelly JC, Joseph AP, Srinivasan N, de Brevern AG (2011) iPBA: a tool for protein structure comparison using sequence alignment strategies. *Nucleic Acids Res* 39:W18–W23
56. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34:W116–W118
57. Liang J, Edelsbrunner H, Woodward C (1995) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci* 7:1884–1897
58. Ikai AJ (1980) Thermostability and aliphatic index of globular proteins. *J Biochem* 88:1895–1898
59. Guruprasad K, Reddy BV, Pandit MW (1990) Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Prot Eng* 4:155–164
60. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132
61. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208
62. Uversky VN, Oldfield CJ, Dunker AK (2005) Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 18:343–384
63. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) Intrinsic disorder and protein function. *Biochemistry* 41:6573–6582
64. Dyson HJ, Wright PE (2002) Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 12:54–60
65. Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z (2001) Intrinsically disordered protein. *J Mol Graph Model* 19:26–59

66. Argos P (1990) An investigation of oligopeptides linking domains in protein tertiary structure and possible candidates for general gene fusion. *J Mol Biol* 211:943–958
67. Melo F, Feytmans E (1998) Assessing protein structures with a non-local atomic interaction energy. *J Mol Biol* 277:1141–1152
68. Bodade RG, Beedkar SD, Manwar AV, Khobragade CN (2010) Homology modeling and docking study of xanthine oxidase of *Arthrobacter* sp. XL26. *Int J Biol Macromol* 2:298–303
69. Ginalski K (2006) Comparative modeling for protein structure prediction. *Curr Opin Struct Biol* 16:172–177
70. Rhodes G (2006) Crystallography made crystal clear, 3rd edn. Academic, Burlington, p 261
71. Heinig M, Frishman D (2004) STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res* 32:W500–W502
72. Ubhayasekera W, Tang CM, Ho SW, Berglund G, Bergfors T, Chye ML, Mowbray SL (2007) Crystal structures of a family 19 chitinase from *Brassica juncea* show flexibility of binding cleft loops. *FEBS J* 274:3695–3703
73. Ubhayasekera W (2011) Structure and function of chitinases from glycoside hydrolase family 19. *Polym Int* 60:890–896
74. Ubhayasekera W, Rawat R, Ho SW, Wiweger M, Von Arnold S, Chye ML, Mowbray SL (2009) The first crystal structures of a family 19 class IV chitinase: the enzyme from Norway spruce. *Plant Mol Biol* 71:277–289
75. Mizuno R, Fukamizo T, Sugiyama S, Nishizawa Y, Kezuka Y, Nonaka T, Suzuki K, Watanabe T (2008) Role of the loop structure of the catalytic domain in rice class I chitinase. *J Biochem* 143:487–495
76. Fukamizo T, Koga D, Goto S (1995) Comparative biochemistry of chitinases-anomeric form of the reaction products. *Biosci Biotech Biochem* 59:311–313