ISSN 1996-3351

Asian Journal of Biological Sciences

http://knowledgiascientific.com

Knowledgia SCIENTIFIC
A Place to Publish Outstanding Research

Asian Journal of Biological Sciences 7 (3): 76-110, 2014
ISSN 1996-3351 / DOI: 10.3923/ajbs.2014.76.110

Cloning, Homology Modeling and Active Site Prediction of Secreted Serine Proteinase (PrDI) and Carboxylesterase (CaDI) Gene from Nematophagous Fungi *Dactylellina cionopaga*

Hanying Yu
Department of Petroleum Engineering, Northeast Petroleum University, Daqing, China

ABSTRACT

Nematophagous fungi are important biocontrol agents of nematode pests. Two genes encoding a secreted serine proteinase (PrDI) and a carboxylesterase (CaDI), respectively, were cloned and characterized by phylogeny and structural comparison. The genes presented the enhanced expression when *Dactylellina cionpaga* was cultured on water agar containing *Caenorhabditis elegans* in contrast to that without the nematodes. Ser225, Glu358 and His470 consisted of the active site of CaDI according to the alignment. The comparison of the structure and its active sites of CaDI with other lipases and carboxylesterases indicated that the protein was more similar to the carboxylesterase and lipases of fungal pathogen of pests, might be butyrylcholinesterase and acetylcholinesterase and probably catalyzed the long-chain carboxylester. The catalytic region of PrDI was identified by docking of the heptapeptide inhibitor Pro-Ala-Pro-Phe-Ala-Ser-Ala. Comparison of the substrate-binding sites of PrDI with those of PR1, Ver112 and VCP1, the serine proteinases from nematophagous fungi and entomopathogenic fungus, showed that the catalytic regions among these serine proteinases were conserved but diverse.

Key words: *Dactylellina cionpaga*, serine proteinase, carboxylesterase, homology modeling, docking

INTRODUCTION

Biocontrol agents provide an environmentally friendly and effective management of plant parasitic nematodes. Nematophagous fungi, for example, *Arthrobotrys irregularis* (Cayrol, 1983), *A. robusta* (Cayrol *et al*., 1978) and *Paecilomyces lilacinus* (Davide and Zorilla, 1983) have been developed as commercial products. However, these bio-pesticides cannot control nematodes quickly now. The mechanisms by which nematophagous fungi infect nematodes have been the basis for development of genetically engineered strains which are able to control parasitic nematodes more effectively than wild types. During the process of penetration by nematophagous fungi, the roles of secreted hydrolytic enzymes are very important.

The proteinases of nematophagous fungi which digest the cuticle or egg shell, are the characteristics of virulence. The cuticle and egg shell of nematodes activates the expression of the enzymes (Tunlid *et al*., 1994; Ahman *et al*., 1996). At least 16 proteinases of nematophagous fungi, all of which were serine proteinases, homologous to each other and belong to proteinase K family (Peptidase_S8, Subtilase) (Tunlid *et al*., 1994; Segers *et al*., 1995; Bonants *et al*., 1995), were purified and characterized (Huang *et al*., 2004; Yang *et al*., 2007a). Recently, an alkaline serine proteinase Hasp and a neutral serine proteinase Hnsp from the endoparasite

Hirsutella rhossiliensis, Dv1 from the nematode-trapping fungus *Dactylellina varietas*, Lmz1 from the toxic fungus *Clonostachys rosea* and an extracellular serine protease Ac1 from the nematode trapping fungus *A. conoides* have been analyzed (Wang *et al*., 2007, 2009; Zhao *et al*., 2005; Yang *et al*., 2007b, c). Hasp and Hnsp, with molecular weights of 33 and 32 kDa, optimal temperatures of 75 and 40°C and optimal pH of 9 and 7, respectively, digested cuticle proteins *in vitro* (Wang *et al*., 2007, 2009). Lmz1 showed a molecular mass of approximately 33 kDa, pI 10.5, optimal activity of Lmz1 at 60°C and pH 11-12. The nematotoxic activity of Lmz1 involving in fungal infection was confirmed by nematode-immobilisation, cuticle degradation and immunodepletion (Zhao *et al*., 2005). Ac1 has a molecular mass of 35 kDa, optimum activity at pH 7.0 and 53.2°C. The protein can degrade not only casein, gelatin, bovine serum albumin, collagen and nematode cuticles, but also immobilize the free living nematode *Panagrellus redivivus* and the pine wood nematode *Bursaphelenchus xylophilus* (Yang *et al*., 2007b). Dv1 of *D. varietas* could degrade casein, gelatin, BSA (bovine serum albumin), hydrolyze the purified cuticle of *P. redivivus* and immobilize the free-living nematodes *P. redivivus* and *Caenorhabditis elegans* (Yang *et al*., 2007c).

Structures of serine proteases of nematophagous fungi are important for explaining host preference of these fungi. Liu *et al*. (2007) modeled the structures of PR1 from entomophagous fungus, Ver112 and VCP1 from nematophagous fungi and suggested the difference of these serine proteinases in the electrostatic surface potential, hydrophobicity and size of the S4 substrate-binding pocket and the number and distribution of hydrogen bonds and salt bridges within regions that are part of or in close proximity to the S2-loop. Crystal structures of PL646, a cuticle-degrading serine protease from nematophagous fungus *P. lilacinus* and Ver112 were reported and the differences between the structures were residues of the substrate binding sites S1 and S4 (Liang *et al*., 2010). Their surfaces likely play important roles during fungal infection against nematodes (Liang *et al*., 2010). P32 structure of nematode eggs parasites *Pochonia rubescens* was predicted (Liang *et al*., 2011). The catalytic pocket of P32 showed differences in the amino acids of the substrate-recognition sites compared with the catalytic pockets of Pr1A and VCP1 proteases (Larriba *et al*., 2012).

Apart from major virulence factor serine proteinases, many other proteins play a role in the infection of the nematode by nematophagous fungi, such as chitinase and collagenase (Yang *et al*., 2007a). However, their functions have not been studied in detail. Ahren *et al*. (2005) compared the global patterns of gene expression in traps and mycelia of the fungus *Monacrosporium haptotylum* which produces adhesive knobs. They found that 23.3% (657 of 2822) of the putative genes were differentially expressed in knobs versus mycelia; including the genes involved in regulating morphogenesis and cell polarity, such as rho1, rac1 and ras1 and a rho GDP dissociation inhibitor (rdi1) and the genes involved in stress response, protein synthesis, protein degradation, transcription and carbon metabolism. Proteomic and quantitative PCR (qPCR) analyses of *A. oligospora* revealed that 90 genes were up-regulated at the early stage of trap-formation and most of these genes encoded proteins that were involved in translation, amino acid metabolism, carbohydrate metabolism, cell wall and membrane biogenesis (Yang *et al*., 2011).

The identification of pathogenicity-related enzymes from different nematophagous fungi will help understanding the selection of substrates, host preference and screening the enzymes that can be applied for targeted genetic engineering at the molecular level (Yang *et al*., 2007a). Esters exist in the cell walls of plants and esterase is implicated in the parasitism of phyto-pests (Montes *et al*., 2004). Disrupting *Metarhizium robertsii* Mest1 which exhibited a marked preference for

short-chain fatty acids, reduced virulence and overexpression, increased virulence to caterpillars (*Galleria mellonella* and *M. sexta*), indicating that esterase involved in the infection of animals (Wang *et al*., 2011).

D. cionopaga is a trapping fungus producing an adhesive column which can prey on and infect phytoparasitic nematodes such as *Meloidogyne javanica* and *Heterodera schachtii*. There are few studies of proteinases and other parasitism-related proteins of adhesive column-developing fungi. A model system involving *D. cionopaga* and *Canenorhabditis elegans* to study the genes involved in preying and infection of nematodes, was used. A cDNA library of this fungus was constructed and those genes coding for the secreted proteins were sequenced through signal peptide trapping (Yu *et al*., 2011). CaDI encoding a carboxylesterase and PrDI encoding a secreted serine proteinase were cloned and were characterized by sequence analysis, quantitative RT-PCR and structural modeling.

MATERIALS AND METHODS

Culturing and RNA extraction: *D. cionopaga* was cultured in PDB for 3d and then collected by filtering through four layers of gauze. After washing in sterilized water three times, the filaments were evenly spread on glass paper on water agar. Live *C. elegans* cultured on oats were added to the hyphae every day, to encourage the formation of trapping columns (Fig. 1). RNA was sequentially extracted by RNAtrip (Applygen Technologies Inc.) from filaments cultured for 2, 3, 4 and 5 day at 25°C. As a control, RNA was also extracted at the same time from the fungus grown on water agar without nematodes. The total RNA extracted was treated with Deoxyribonuclease I (Life Technologies Corporation) and reverse transcribed to synthesize the first strand cDNA using PrimeScript™ RT Enzyme Mix I with the Oligo dT Primer (50 μM) and Random 6 m (100 μM), by incubation at 37°C for 30 min.

Cloning and sequence analyses of CaDI and PrDI: The Open Reading Frame (ORF) and the 5'-terminal fragment of CaDI were amplified using inverse PCR. The genomic DNA of *D. cionopaga*

Fig. 1: *D. cionopaga* infect the nematode *C. elegans*

was digested by SacI and SacII, respectively and purified using a Cycle-Pure Kit (Omega Bio-Tek, Inc.). Self-circularization of the digested and purified DNA was performed with T4 DNA ligase. Using the intramolecular ligated genomic DNA digested by SacI as the templates, two primer pairs, Er6/Er7 and Er1/Er2, designed according to the known sequences of CaDI, found by signal peptide trapping and sequencing (Yu *et al*., 2011), were, respectively used for the nest PCR with ExTaq in the touchdown reaction conditions. When using the genomic DNA digested by SacII as the templates, the nest PCR was carried out using the primer pairs Er6/Er1 and Er6/Er7, respectively, to obtain the complete open reading frame of CaDI. The expected fragments were purified through agarose gel, using the PCR products with single primer as the negative control. The cDNA of CaDI was also amplified by nest PCR using the primer pairs Er-L/Er-R1 and Er-L/Er-R, respectively, from the total RNA of *D. cionopaga* induced by nematodes and subsequently inserted into pMD18-T and sequenced.

The genomic DNA sequences of PrDI were amplified with the primers Pn-S-1 and Pn-S-2 at the annealing temperature of 57°C, based on the known cDNA sequence found by signal peptide trapping and sequencing (Yu *et al*., 2011). The sequencing of the obtained gene fragments was accomplished by Sangon Biotech (Shanghai) Co., Ltd.

The probes for PrDI and CaDI were labeled, respectively, with the templates of the full-length ORFs. Genomic DNA of *D. cionopaga* extracted using CTAB solution was digested by SalI, or XhoI. A Southern blot analysis was performed following the manufacturer's methods of the DIG High Prime DNA Labeling and Detection Starter Kit I (F. Hoffmann-La Roche Ltd.).

The sequences of PrDI and CaDI were submitted to the online programs BLASTN and BLASTX to enable hypotheses about their functions. Possible regulation sites and introns were analyzed. The putative protein sequences were submitted to BLASTP and Motifscan to find homologs and motifs. Phylogenetic trees were constructed using the MEGA program (Thompson *et al*., 1994).

Real-time fluorescent quantitative RT-PCR: The first strand cDNA synthesized as mentioned above was used as templates for quantitative RT-PCR with an *in vitro* control. The primers BT2a and BT2b were used to amplify the DNA fragment of the **β**-tublin from *D. cionopaga* (Yang, 2006). The cDNA of the **β**-tublin of the fungus was then amplified using the primers Beta-tublin-L and Beta-tublin-R, cloned into the pMD18-T vector and submitted to SinoGenoMax Co. Ltd., for sequencing. The reference genes (**β**-tubulin gene and 18s rRNA gene) were evaluated by the GeNorm algorithm. The amplification of the **β**-tubulin gene showed that it was transcribed uniformly in the filaments cultured on the WA (water agar) and WA+CE (water agar with *C. elegans*) plates ($p<0.05$) and could be used as an external reference (Fig. 2).

The primers specific to **β**-tublin, PrDI and CaDI of *D. cionopaga* were designed for quantitative PCR analysis by the software Primer Express 2.0. Beta-tublin-L and Beta-tublin-R were used for the **β**-tublin amplification, P-r-L and P-r-R for PrDI and C-r-L and C-r-R for CaDI which are listed in Table 1.

The qRT-PCR was performed using the kit for the Bio-Rad iQ5 Gradient Real Time PCR Cycler and the reaction conditions were 95°C for 10 sec, 95°C for 5 sec, 60°C for 30 sec, for 40 cycles on a Bio-Rad iQ5 cycler. Another series of cycles (95°C for 1 min, 55°C for 1 min, 55°C for 30 sec, for 81 cycles, temperature change 0.5°C) were suspended to draw the melting curves. The melting curves indicated that these primers could, respectively amplify the unique genes in the conditions described. The expressions of PrDI and CaDI from *D. cionopaga,* whether induced by nematodes or not, were homogenized to that of **β**-tublin by the relative standard curve method. The relative quantitative PCR was repeated three times and every treatment included three replicates.

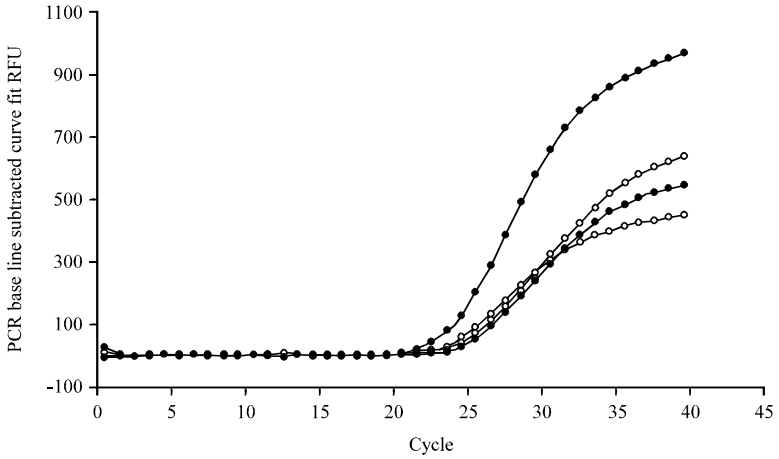

Fig. 2: Amplification of the β-tubulin gene (in duplicate) with the total RNA extracted from the filaments cultured on WA or WA+CE as the templates

Table 1: Primers used in this study

Primers	Sequences
Er6	5'-TAGACAATCTTCGTTCCC-3'
Er7	5'-TTAGTTTGGATCCACGGT-3'
Er1	5'-GATAATCAAGGAGAACGGTAACAC-3'
Er2	5'-AGCCCAGATCAACTTGAGGAG-3'
Er-L	5'-ATGCGTCTTACGACAATAATCTTTGCC-3'
Er-R1	5'-AGAGGGCATGTCACAACACTTAGAT-3'
Er-R	5'-TTAGATAAAGTAGGCGTATTGAGGACC-3'
C-r-L	5'-GAAATGTGACGAGCGGGTTG-3'
C-r-R	5'-TTGGAGGGACAGCGTAGCG-3'
Pn-S-1	5'-TAGCGGCCGCTAATGTTTTCCAAGAGCCTCATCAC-3'
Pn-S-2	5'-ATAGGGCCCTTTAAGTGGAGCAAGTCCTGAGGGA-3'
P-r-L	5'-TGCCGCTCTTACCGCTGGTG-3'
P-r-R	5'-GCGTCGTCCGATTCGTTT-3'
BT2a	5'-GGTAACCAAATCGGTGCTGCTTTC-3'
BT2b	5'-ACCCTCAGTGTAGTG ACCCTTGGC-3'
Beta-tublin-L	5'-TCGGTGCTGCTTTCTGGCAAAC-3'
Beta-tublin-R	5'-GCACGTGGAACATATTTGTTGCCC-3'

Physicochemical property calculations: The isoelectric point (pI), extinction coefficient at 280 nm and the hydrophobicities of mature proteins were calculated, respectively using the online tools ProtParam and ProtScale (Gill and von Hipple, 1989; Gasteiger *et al.*, 2005). The location in the cells of the proteins was predicted by Signal P (Petersen *et al.*, 2011). The molecular weight, amino acid composition and charge distribution of each mature enzyme was calculated using the program SAPS (Brendel *et al.*, 1992).

Homology model: Homology modeling, structural optimism by energy minimization using GROMACS and assessment by PROCHECK were performed following the methods of (Liu *et al.*, 2007). The modelings were performed using the MODELLER 9.10 package. Twenty models were generated from the templates and were clustered. An optimized average model was generated from

these models. The structure of CaDI was modeled using the template of the EstA (Protein Data Bank (PDB) accession code 1UKC) of *Aspergillus niger* which has the similarity of 47% to the target protein (Bourne *et al*., 2004). The putative carboxylesterase (signed as CaA1) of the pathogenic and allergenic *A. fumigatus* (GenBank accession No. XP_752911) was modeled with the templates of 3AINA which was a thermostable carboxylesterase from *Sulfolobus tokodaii* (Angkawidjaja *et al*., 2012) to present the differences of CaDI from carboxylesterase of pathogens. The similarity of carboxylesterase of *A. fumigatus* with 3AINA was 26%. The carboxylesterase of the human pathogen *Cryptococcus neoformans* (GenBank accession No. XP_568314) was also modeled, signed CaC4, with EstA protein, a virulence factor from *Streptococcus pneumoniae* (PDB accession code 2UZ0A) (Kim *et al*., 2008) as the template with 24% similarity. The X-ray structure of the serine protease proteinase K (PDB accession code 1IC6) (Betzel *et al*., 2001) from *Tritirachium album limber* was applied to model PrDI. The similarity between the sequence of PrDI and that of 1IC6 which did not have a signal peptide and propeptide, was 43%. The developed structure was signed as PrDI_1. Besides 1IC6, the structure of PrDI was also modeled using 1P7WA as the templates which had the same sequence with 1IC6 and was crystallized with an inhibitor of N-Pro-Ala-Pro-Phe-Ala-Ser-Ala-C. The developed structures were signed as PrDI_i. PR1 produced by an entomopathogenic fungus *Metarhizium anisopliae* which is an insect cuticle-degrading serine protease (St. Leger *et al*., 1987, 1992), VCP1 from the nematophagous fungus *Pochonia chlamydosporia* which played an important role during the infection process of nematode eggs (Segers *et al*., 1996; Morton *et al*., 2003) and Ver112 from the nematophagous fungus *Lecanicillium psalliotae* (Yang *et al*., 2005a, b) which exhibited cuticle-degrading and nematicidal activity, were modeled with 1P7WA as the template which were signed as PR1_i, VCP1_i and Ver112_i. The similarity of PR1, VCP1 and Ver112 with 1IC6 was 68, 64 and 66%, respectively. The models described here were referred to as the PDB code, except that were mentioned. The amino acid numbering started from the first amino acid of the proteins.

Structural analyses and comparisons: The number of hydrogen bonds and radii of gyration of the proteins PrDI and CaDI were calculated using GROMACS software (g_gyrate and g_hbond) (Lindahl *et al*., 2001). The numbers of salt bridges of the proteins were computed using the online program ESBRI (Kumar and Nussinov, 1999, 2002; Kumar *et al*., 2000; Sarakatsannis and Duan, 2005). The active site residues of the target enzymes were predicted based on the motifs and alignment with the homologues. The Root Mean Square Deviation (RMSD) of PrDI and CaDI with the templates and the homologues was calculated using Swiss-pdb Viewer software (Guex and Peitsch, 1997). Secondary structure element and Solvent-accessible Surface Area (SASA) of the models were analyzed using the DSSP program (Kabsch and Sander, 1983).

Automated docking: Heptapeptide inhibitor (Pro-Ala-Pro-Phe-Ala-Ser-Ala, PDB code: 1P7WB) was chosen to dock with the proteinases to analyze the catalysis region. Polar hydrogen atoms were added to the molecules and the flexible residues of the acceptors were assigned according to the motif of active sites. Automated docking was carried out using the Lamarckian genetic algorithm and AutoDock Tools (ADT) (Morris *et al*., 2009). The appropriate binding orientations and conformations of the inhibitor in the binding pocket were selected (Huey *et al*., 2007). Calculations of molecular surface and electrostatic potential of active sites were performed using the program Autodock v. 4.2. The binding affinities were predicted by Autodock (Adinarayana and Devi, 2011).

RESULTS AND DISCUSSION

Carboxylesterase

Cloning, phylogeny and expression analyses of CaDI: CaDI of *D. cionopaga* was cloned by inverse PCR. The open-reading frame of the gene, with one copy in the genome (Fig. 3) was 1833 bp in length and encoded 610 amino acids. The sequence comparison of cDNA and genomic DNA of CaDI (GenBank accession No. JF699728.1) showed that six introns exist in the genomic DNA sequence of this gene and that all conformed to the GT-AG rule of extron-intron splicing. CaDI contained the esterase-lipase motif, carboxylesterases type-B serine active site and signature 2 and the specific motif serine-rich region profile and copper binding octapeptide repeat motif, suggesting that probably the copper ions were bound to the protein *in vivo*. The phylogeny of CaDI and some proteins in the carboxylesterase/lipase superfamily of fungi is shown in Fig. 4. Quantitative RT-PCR was carried out to analyze the differential expression of CaDI to present the function of the gene in the parasitism of *D. cionopaga*. The result of quantitative RT-PCR indicated that CaDI was expressed 4-fold higher than the control when the mycelia were inoculated with nematodes for 2 day (Fig. 5). The differential transcription of this gene decreased faster with inoculation time and became even lower than the control, suggesting that CaDI played a role in early infection of the nematodes and might be involved in resistance to adverse circumstances or in the self-lysis of *D. cionopaga*. However, no regulation sites such as AreA, CreA and PacC, usually implicated in the expression of parasitism genes like the serine proteinase gene of *M. megalosporum* and were found upstream of CaDI (Kanda *et al*., 2008).

Physicochemical properties of CaDI: The pI and molecular mass of the putative protein encoded by CaDI of *D. cionopaga* were 8.65 and 64.5 kDa and those of the mature proteins were 8.51 and 62.7 kDa, respectively (Table 2). The signal peptide (MRLTTIIFAALAVAQA) suggested

Table 2: Calculated physicochemical properties of PrDI and CaDI

Protein	Molecular weight (kDa)	pI	PPNN[a]	Extinction coefficient (M^{-1} cm^{-1}, 280 nm)
PrDI	33.3	5.40	21/26	50880
CaDI	62.7	8.51	41/35	86220

[a]Proportion of positively charged and negatively charged residue number

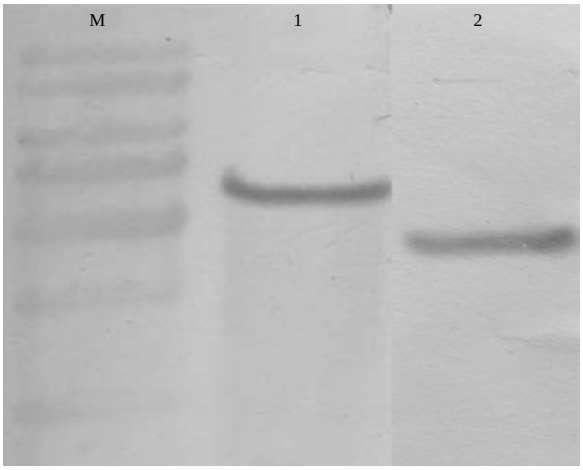

Fig. 3: Southern hybridization. Lane M: Molecular weight marker; Lane 1: CaDI, genome digested by SalI; Lane 2: PrDI, genome digested by XhoI

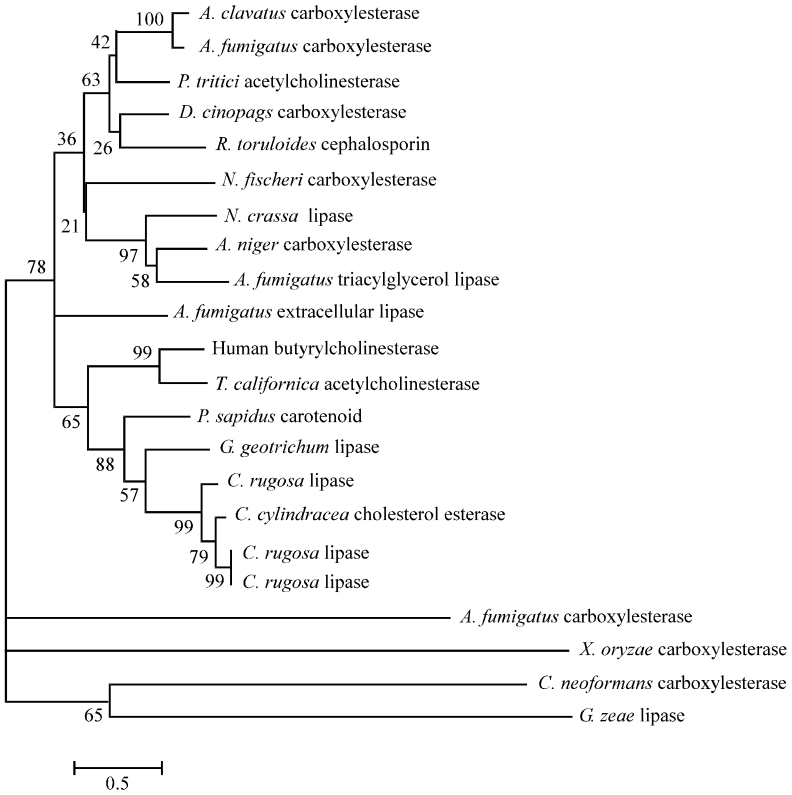

Fig. 4: Phylogeny of CaDI with the program MEGA. *C. neoformans* carboxylesterase, XP_568314; *A. fumigatus* carboxylesterase, XP_752911; *X. oryzae* carboxylesterase, YP_199165, 3H2G; *D. cionopaga* carboxylesterase, AEN99944.1; *A. fumigatus* extracellular lipase XP_746814.1; *A. niger* carboxylesterase, AAS13488, 1UKC; *N. fischeri* carboxylesterase, XP_001260357.1; *R. toruloides* cephalosporin, AAB93483.1; *A. clavatus* carboxylesterase, XP_001276049; *A. fumigatus* carboxylesterase, XP_753283.1; *P. tritici* acetylcholinesterase, XP_001931700.1; *P. sapidus* carotenoid, CAH17527.1; *G. geotrichum* lipase, ABC39650.1; *A. fumigatus* triacylglycerol lipase, XP_747228.1; *N. crassa* lipase, CAE76488; Human butyrylcholinesterase, 1P0Q; *T. californica* acetylcholinesterase, 1HBJ; *C. rugosa* lipase, 1LPP, 1CRL, 1GZ7; *C. cylindracea* cholesterol esterase, 1LLF; *G. zeae*, 3NGM

that CaDI is an extracellular protein. The extinction coefficient of the mature protein CaDI at 280 nm was 86,220 M^{-1} cm^{-1}. The numbers of the positively and negatively charged residues were 41 and 35. The putative Asn-linked glycosylation sites, amidation site, cAMP-and cGMP-dependent protein kinase phosphorylation site, casein kinase II phosphorylation site, N-myristoylation site, protein kinase-C-phosphorylation site and tyrosine kinase phosphorylation site were identified (Table S1, Supplemental data), implying that the protein was glycosylated, amidated and phosphorylated. These sites were not found based on the modeled structure by PDB-site-scan. However, the glycosylation sites of human butyrylcholinesterase (PDB accession code 1P0Q) and the esterase of *A. niger* (1UKC) also could not be found by PDB site scan, although they existed in the proteins. Bourne *et al.* (2004) proposed that the glycans contributed to an active role in EstA biological function, perhaps in recognizing a yet unidentified partner (Bourne *et al.*, 2004). Six Asn-linked glycosylation sites were found in EstA. The analysis of the post-translational

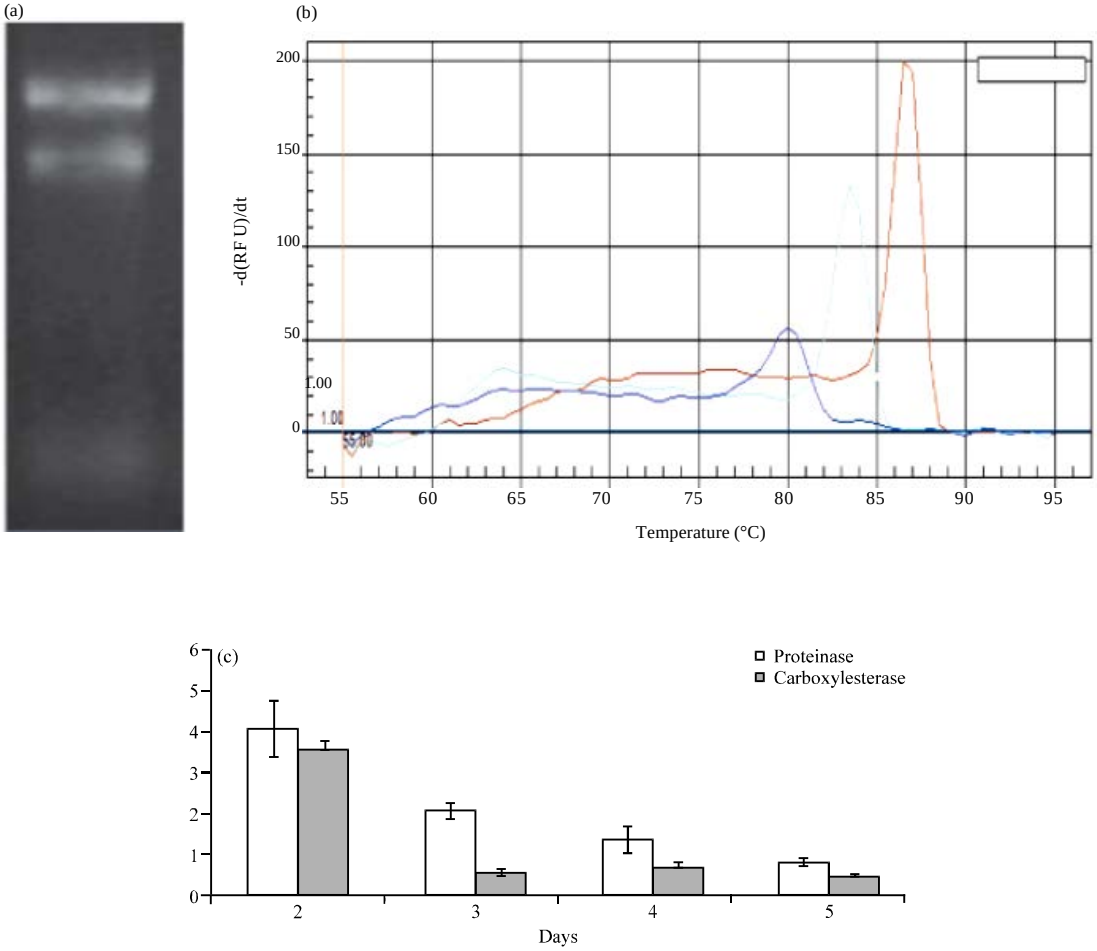

Fig. 5(a-c): Quantitative RT-PCR of PrDI and CaDI, (a) Total RNA of *D. cionopaga*, (b) Melting curves and (c) Differential transcription fold. Error bars indicate standard deviations from three independent experiments

modification site of CaDI indicated that three Asn-linked glycosylation sites which were not conserved and existed in the protein. The sites were located on the same side as the active sites that the protein opened in. The active sites of CaDI were determined to be Ser225, Glu358 and His470 according to alignment with other carboxylesterases. The posttranslation sites of the protein were mapped on its modeled structure. Asn-linked glycosylation site (Asn100-Ala103), casein kinase II phosphorylation site (Thr515-Glu518), protein kinase-C-phosphorylation site (Thr515-Arg517) and N-myristoylation sites (Gly142-Phe147, Gly223-Gly228, Gly316-Trp321, Gly482-Leu487) were in close proximity to the active sites, suggesting the importance of these sites in the function of the active sites.

Model description and structural comparison: A Ramachandran plot analysis showed 12 residues of CaDI, 3 residues in the active site pocket, 3 residues in the protein fragment (104-476) and no active sites located in the disallowed region (Fig. 6). The constructed model of CaDI consisted of 15 β strands and 18 α helixes and had the α/β hydrolase fold in the

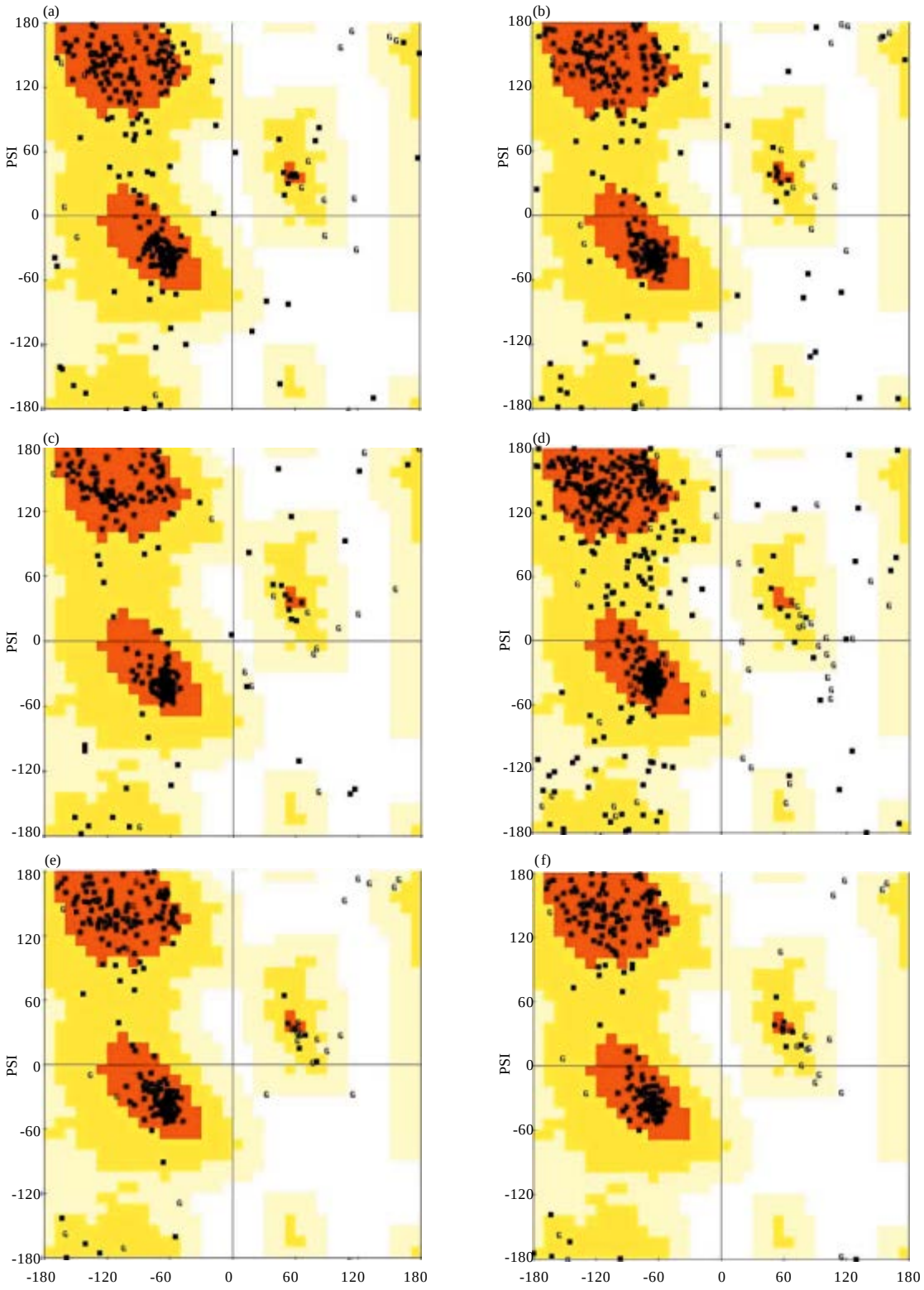

Fig. 6(a-h): Continue

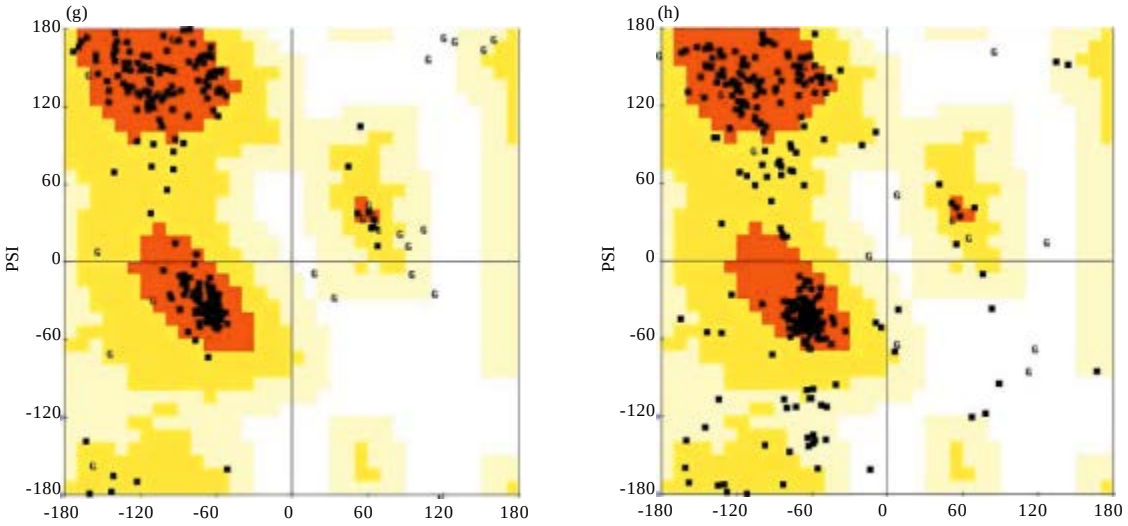

Fig. 6(a-h): Ramchandran plots of the constructed models, (a) PrDI_i, (b) PrD_1, (c) CaC4, (d) CaDI, (e) VCP1_i, (f) Ver112_i, (g) PR1_i and (h) CaA1

Table 3: Geometrical properties of PrDI and CaDI

Proteins	RMSD (Å)[a]				SASA (Å²)[b]	Rg (Å)[c]	NHB[d]	NSB[e]	α helix[f]	β sheet[g]	β turn[h]
PrDI_i	1P7W	Pr1_i	VCP1_i	Ver112_i	13129.7	1.77	201	105	70	77	171
	0.49	2.53	5.07	1.60							
PrDI_1	1IC6	PR1	VCP1	Ver112	13158.1	1.77	201	126	76	70	172
	0.32	0.41	0.46	0.39							
CaDI	1UKC	3NGM	CaA1	CaC4	21930.1	2.29	298	187	156	93	340
	0.65	1.37	2.11	2.31							
	1CRL	1LLF	1LPP	1HBJ	1P0Q						
	1.42	1.36	1.39	1.45	1.40						

[a]RMSD, backbone RMSD with respect to PrDI or CaDI. [b]SASA, Total solvent accessible surface area, [c]Rg, Radius of gyration, [d]NHB, No. of hydrogen bonds (The hydrogen-donor-acceptor angle is equal to or smaller than 30° and the donor-acceptor distance is equal to or smaller than 0.35 nm), [e]NSB, Number of salt bridges, A possible salt bridge is considered to exist if the distance between two oppositely charged residues is within 6 Å, [f]No. of amino acids in the α helix, [g]No. of amino acids in the β sheet, [h]No. of amino acids in the β turn

esterase-lipase superfamily: A 12-stranded central β strand surrounded by 13 α helices (Fig. 7a) (De Simone *et al*., 2004). The Root-mean-square Deviation (RMSD) between the model and the template was 0.65Å. The solvent-accessible surface area of the model was 21930.1Å^2 (Table 3). 298 hydrogen bonds and 187 salt bridges were found to exist in CaDI (Table 3). Alignment with the template showed that the three possible disulfide bridges linked the residues 78-120 (anchoring the β4-β5 loop), 281-296 (anchoring α7 helix and α8 helix) and 442-577 (anchoring C-terminal α18 helix and α14 helix). Cys577 that formed an S-S bridge with Cys442 could also be involved in the hydrogen bond network Ser337-[Gln589, Ala584, Arg581, Cys577]. Carboxylesterase could be dimmer or tetramer. The used template of modeling in this study, EstA (1UKC), is a dimmer. However, the results of hydrophobic patch and contact surface presented that no contact surface exsited on CaDI and the enzyme might be a monomer (Fig. 8a).

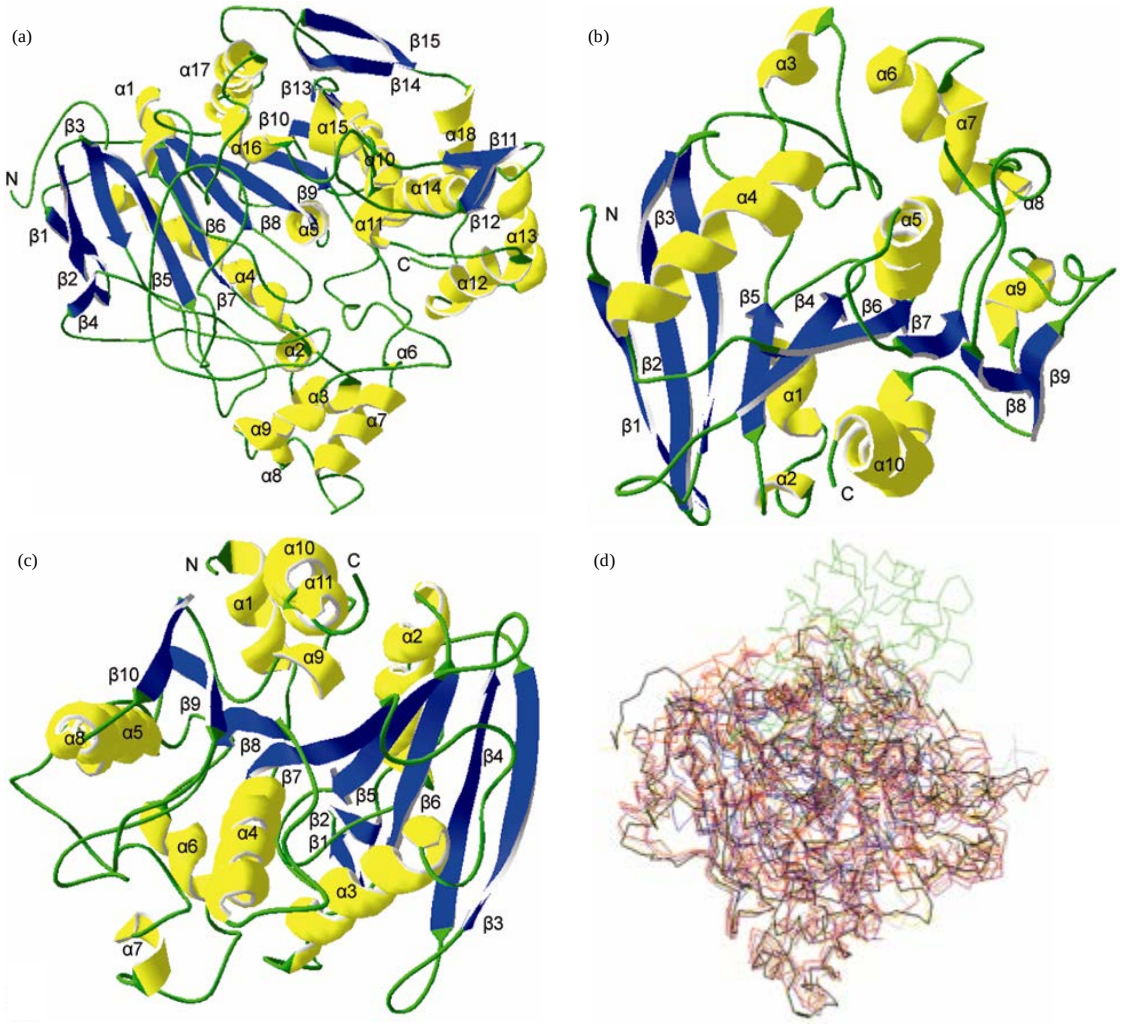

Fig. 7(a-d): Structural comparison of CaDI and the homologs. Ribbon diagrams of (a) CaDI, (b) CaC4, (c) CaA1 and (d) The structural superimposition of CaDI (black), 1UKC (yellow), 3NGM (red), CaC4 (green), CaA1 (blue). 1CRL (purple), 1LLF (pink), 1LPP (orange), 1HBJ (brown) and 1P0Q (grey)

The putative carboxylesterase of *A. fumigatus* and *C. neoformans* which were shown to be related to the pathogenicity of fungi, was modeled to present the differences of CaDI from carboxylesterase of pathogens. The constructed model CaC4 was made up of 9 β strands and 10 α helixes (Fig. 7b) while the model CaA1 constituted 10 β strands and 11 α helixes (Fig. 7c). The Ramachandran plot analysis indicated that 4 disallowed residues were in the CaA1 model while there were 7 residues in the CaC4 model. The molecular surface and electrostatic potential mapped onto the surface are shown in Fig. 8b and c.

The structure of the putative protein of CaDI was compared with other carboxylesterases to analyze the function of the gene further. The Z-scores of the achieved structure of CaDI, calculated using the Dali server, were 54.7 with 1UKC-A, 40.3 with 1P0Q-A (human butyrylcholinesterase)

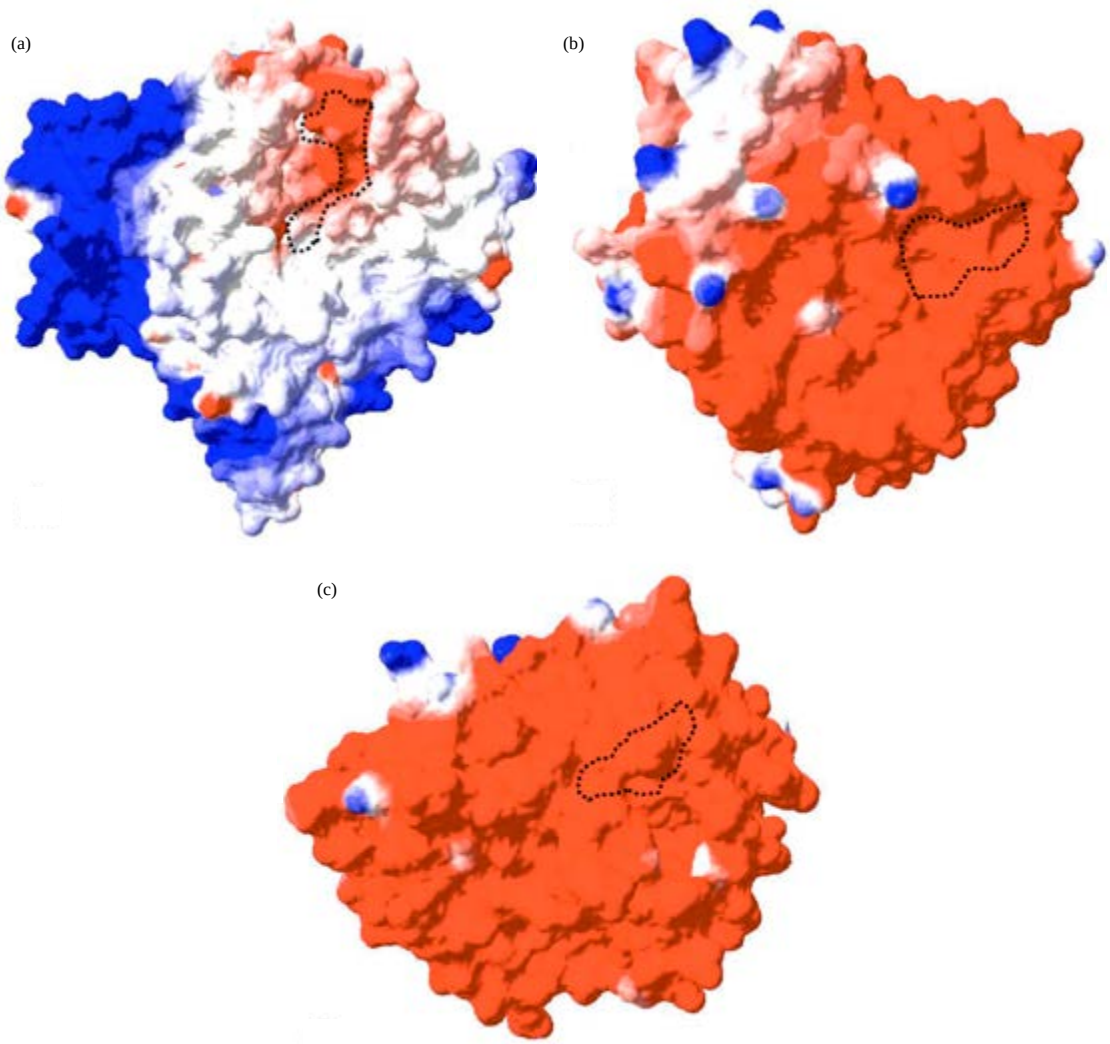

Fig. 8(a-c): Electrostatic surface potential comparison of CaDI and the homologs, (a) CaDI, (b) CaC4 and (c) CaA1. The approximate positions of active sites are dash-dotted

(Nicolet *et al*., 2003), 39.4 with 1HBJ-A (acetylcholinesterase of *Torpedo californica*) (Doucet-Personeni *et al*., 2001) and 38.5 with 1LPP-A (lipase of *Candida rugosa*) (Grochulski *et al*., 1994). The Z-score is the distance, in standard deviations, between the observed alignment RMSD and the mean RMSD for random pairs of the same length, with the same or fewer gaps. Z-scores less than 2 are considered to lack statistical significance. The Z-scores of CaDI with other esterases indicated that the protein was highly similar to that of 1UKC, 1P0Q, 1HBJ and 1LPP which is same with the results of the sequence comparison by E-value. *G. zeae* are pathogens of plant and the lipase of *G. zeae* (3NGM) were reported to be related to parasitism. *C. rugosa* is an emerging human pathogen (Colombo *et al*., 2003). The Z-scores of CaDI with the lipase 1CRL and the cholesterol esterase 1LLF of *C. rugosa*, the constructed models of carboxylesterases CaC4 and CaA1 and 3NGM were computed. The results indicated the Z-scores of CaDI with 1CRL, 1LLF,

CaC4, CaA1 and 3NGM were 37.6, 36.4, 13.8, 17.5 and 4.8, respectively. The RMSD between the structural model of CaDI and the enzymes described above were calculated: 2.31>2.11>1.42>1.37>1.36>0.65 (Table 3). The RMSD and Z-score comparison of structures indicated that the structure of CaDI was more similar to the carboxylesterase and lipases of fungal pathogen of pests than esterases of human pathogens, presenting the similarity to the structures of homologous sequences rather than the relationships of pathogens and enzyme activity. The alignment of CaDI and the other esterases was shown in Fig. 9.

Substrate-binding sites and comparison: The active site Ser225 of CaDI was in β8, the other two were in the β10-α11 and β13-α15 loop, respectively. The residues in the aromatic cluster (Tyr256, Trp321, Phe362, Phe440) formed the acyl/propyl pocket on the surface of CaDI. The active site triad and the aromatic cluster formed the substrate-binding pocket. NH groups of Ala242 and Gly160 formed the oxyanion hole according to the alignment. Charge distribution analysis indicated that Glu358 in the substrate-binding region was a negatively charged amino acid. No charge fragment was related to the active site.

The result of electrostatic potential analysis of CaDI showed a positive potential for the face where the active site cavity opens and a negative potential for the opposite face, leading the substrate into the pocket which was consistent with that of EstA (Bourne *et al*., 2004). However, the negative potential in the structure of CaDI dominated. In contrast, the positive potential was preponderant in CaA1 and CaC4. The active site pockets of CaA1 and CaC4 opened vertical to the electrostatic potential gradient. The active site region of CaDI mainly consisted of hydrophilic amino acids and was negative (Fig. 8) and the protein did not possess a lid like the template EstA which exists in the lipase of *Geotrichum candidum* (Schrag and Cygler, 1993), suggesting that CaDI functions as a carboxylesterase rather than a lipase.

Hydrogen bond calculation indicated that the amino acids in the active site pocket in the secondary structures (β8 and α5, β10 and α11, β13 and α15) beside the active sites and near the active site pocket, are implicated in formation of 52 H-bonds. The active site triad Ser225 was able to form hydrogen bonds with Thr253; similarly Glu358 with Leu487 or Asn355 and His470 with Val474. The hydrogen bond of Thr220-Val136 might anchor β6 and β8 beside the active site Ser225 while Thr253-[Gly223] linked β8 and β9. The hydrogen bond network of Gly240-[Ala237, His234, Ser230, Ala226] linked α5 to the α5-β9 loop. His234/Leu202-[Gln198, Gly194/Ser230] linked α5 and α4. The secondary structure α11 beside the active site Glu358 were stabilized by Val363-Gly359. All the hydrogen bonds involved in the conserved amino acids, indicating the role of these amino acids in the structure of the esterase/lipase. Salt bridge calculation indicated that the amino acids in the active site pocket, in the secondary structures beside the active sites and near the active site pocket, are implicated in formation of 13 salt bridges. Asp197 possibly formed the salt bridges with His234, His350 or Lys460 which could stabilized α5 beside the active site Ser225, β10 and α11 beside the active site Glu358 and β13 and α15 beside the active site His470, presenting the importance of the conserve amino acid. Phe440 in the aromatic cluster formed the hydrogen bond network with Ser444 and Ala436 which was found to stabilize α14. The salt bridge calculation indicated that the four amino acids in the aromatic cluster did not form any ion interactions.

The active sites of CaA1 were in the β6-α5, β8-α9 and β9-α10 loops and the active sites of CaC4 were in the β7, β9-α8 loop and the β10-α9 loops. The conformation differences of the active sites of CaDI, CaA1, CaC4 and other esterases were presented in Table 4.

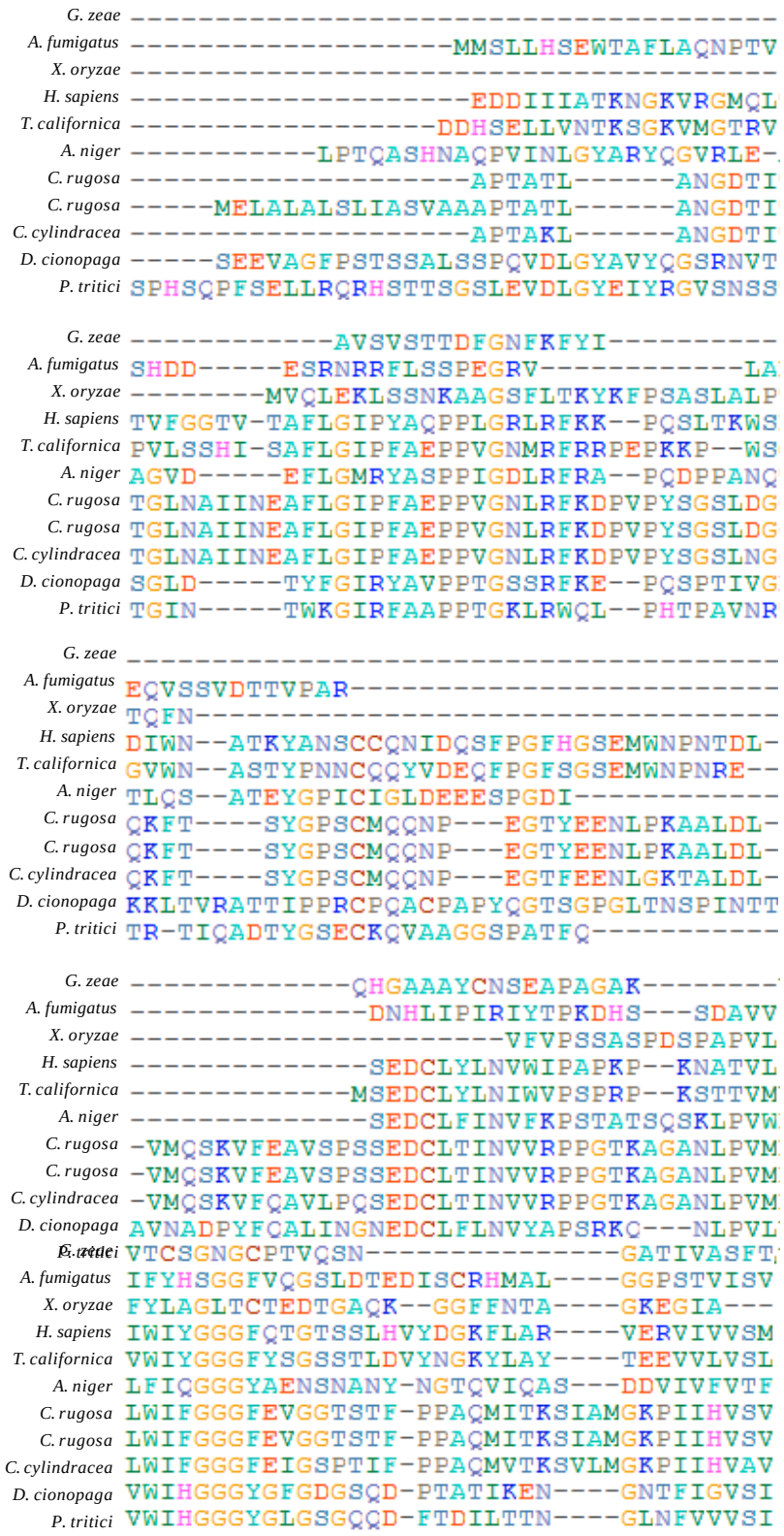

Fig. 9: Continue

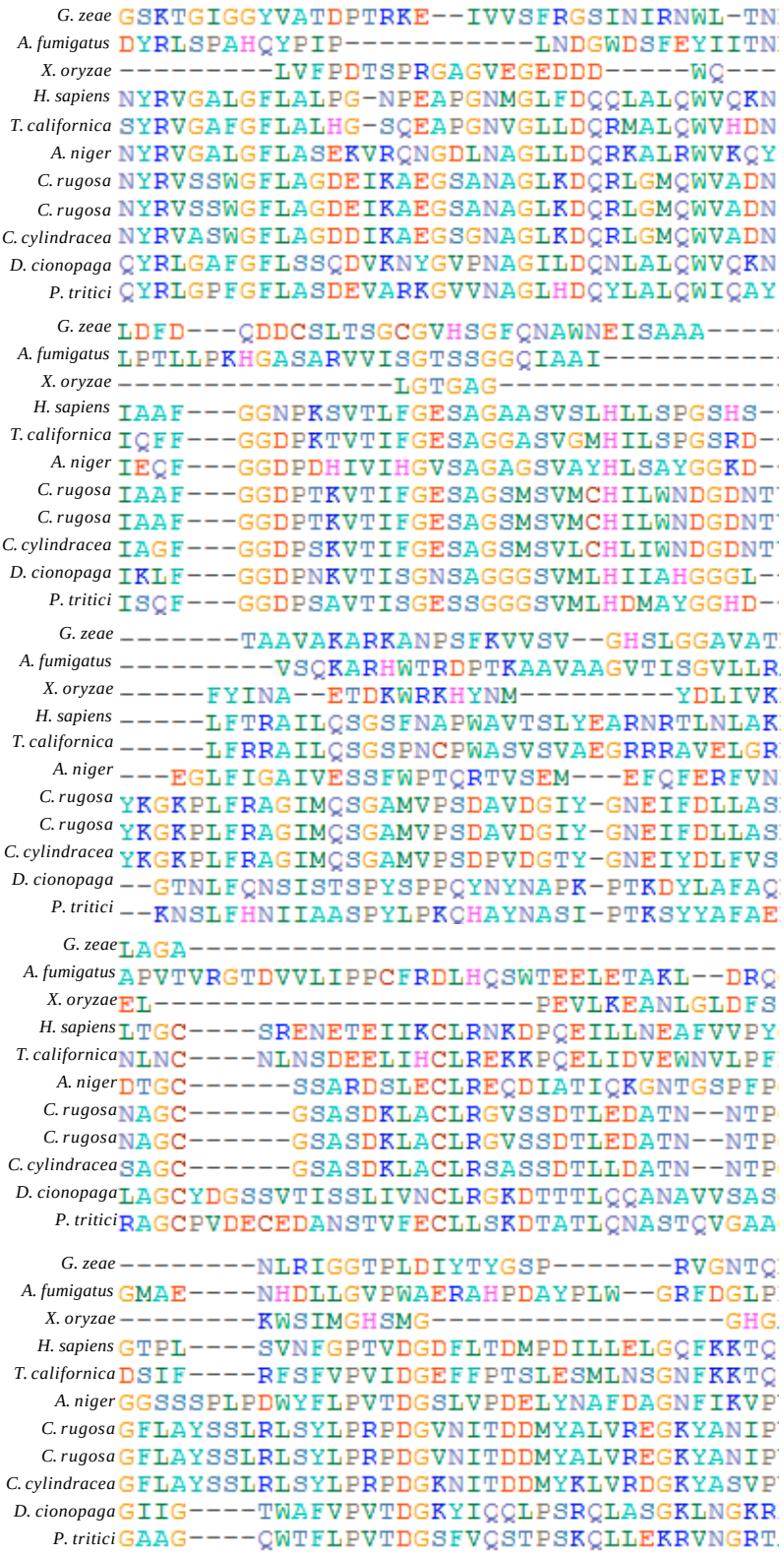

Fig. 9: Continue

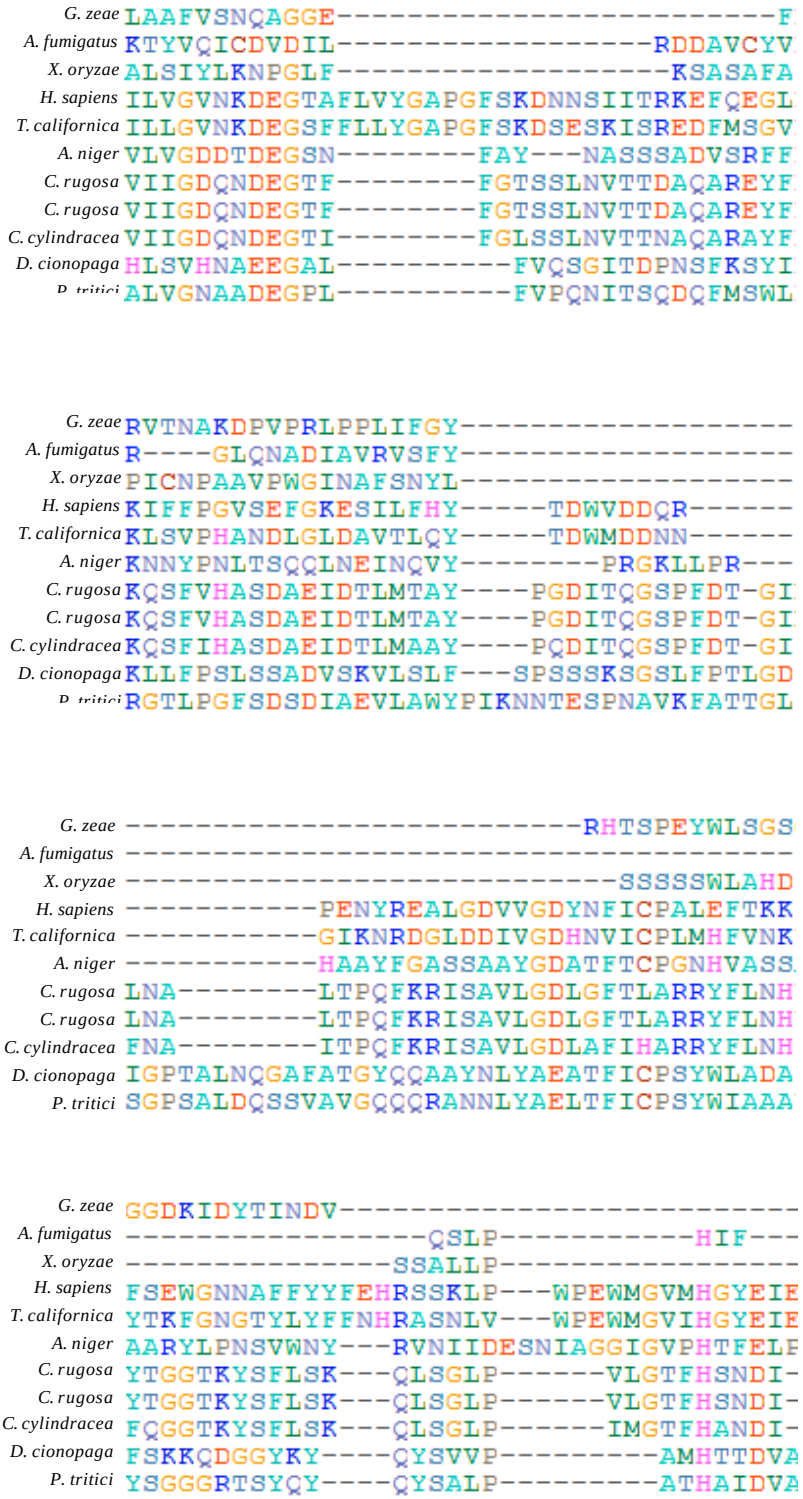

Fig. 9: Continue

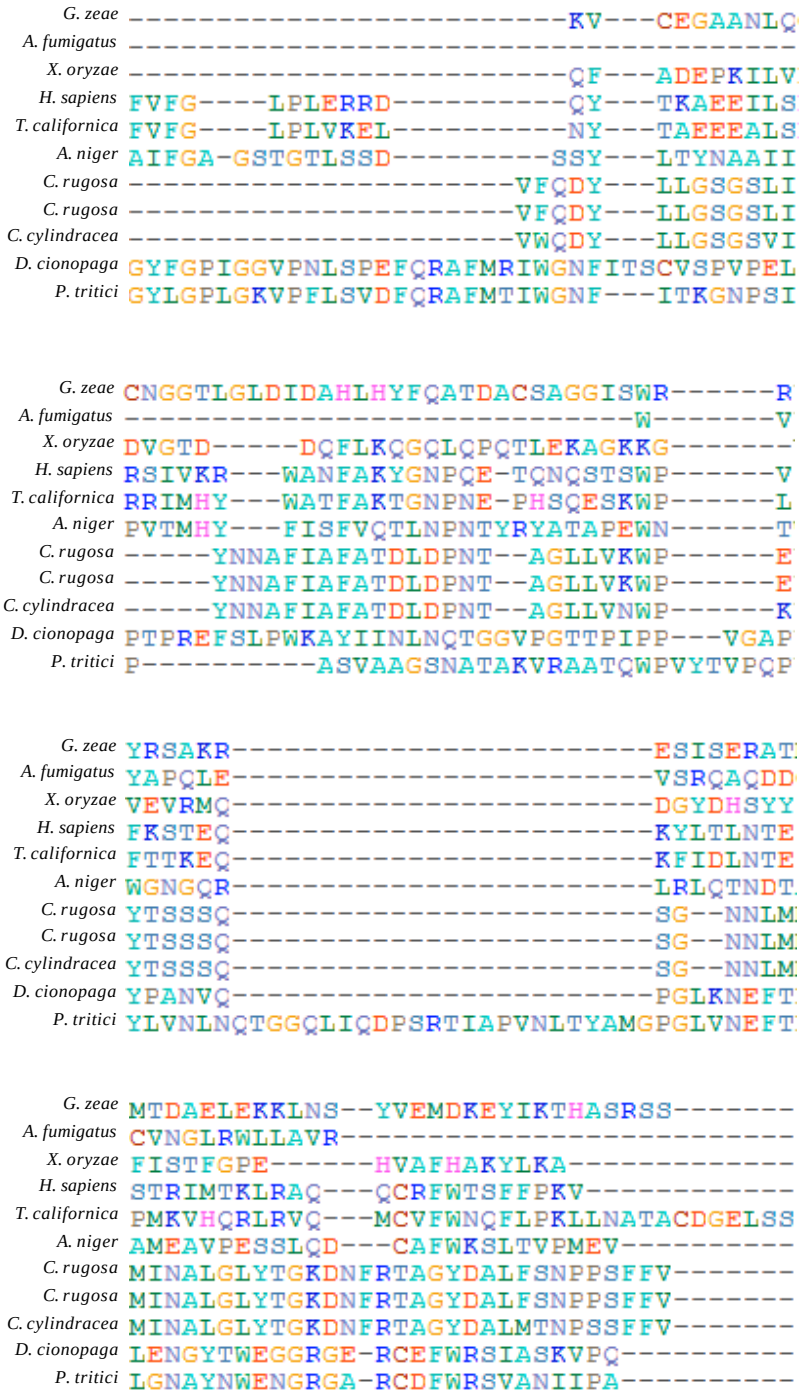

Fig. 9: Alignment of CaDI and other esterases. *G. zeae*, 3NGM; *A. fumigatus*, XP_752911; *X. oryzae*, YP_199165; *Homo sapiens*, 1P0Q; *T. californica*, 1HBJ; *A. niger*, 1UKC; *C. rugosa*, 1LPP, 1CRL; *C. cylindracea*, 1LLF; *D. cionopaga*, AEN99944.1; *P. tritici*, XP_001931700

Table 4: Comparison of the active sites of esterases

Enzyme	Oemga	phi	psi
CaDI			
Ser225	175.339	65.174	-125.419
Glu358	-176.964	-65.723	-39.416
His470	176.767	-82.056	80.463
1UKC			
Ser210	173.839	61.723	-123.013
Glu338	-176.396	-52.414	-38.542
His440	-177.698	-50.345	121.898
CaC4			
Ser152	171.602	62.841	-109.808
Asp227	-174.666	-56.201	116.965
His258	-178.267	-78.337	-22.760
CaA1			
Ser155	169.676	67.353	-119.596
Asp270	-178.388	-145.106	126.170
His300	176.762	-52.142	137.950
1CRL			
Ser209	179.786	56.811	-121.814
Glu341	-177.062	-62.860	-34.274
His449	-178.583	-60.592	126.862
3NGM			
Ser144	179.185	67.240	-149.447
Asp198	176.934	-55.127	124.830
1LLF			
Ser209	176.120	58.433	-114.451
Glu341	-179.587	-50.134	-37.937
His449	178.730	-54.247	129.186
1P0Q			
Ser198	178.813	59.822	-121.957
Glu325	179.615	-49.295	-39.673
His438	177.929	-63.078	127.968
1HBJ			
Ser200	-178.974	57.806	-122.030
Glu327	-179.363	-82.728	1.338
His440	174.275	-45.179	141.344
1LPP			
Ser209	-178.299	49.726	-113.927
Glu341	-178.543	-57.928	-34.608
His449	-177.226	-61.641	139.181

The RMSD of the active site triad between CaDI and the template was 8.12Å. The RMSD comparison of active sites of CaDI, CaA1, CaC4, 1UKC, 3NGM, 1CRL and 1LLF was 32.16>11.96>8.03>7.86>7.83. The structure of active sites of CaDI was more similar to the lipases of fungal human pathogens than esterases. The comparison of active sites of CaDI with 1P0Q-A, 1HBJ-A and 1LPP-A (RMSD: 7.86>7.78>7.74) presented that the conformation of CaDI was more similar to that of butyrylcholinesterase (1P0Q-A). Our phylogeny confirmed that CaDI showed higher similarity with the acetylcholinesterase of *Pyrenophora tritici-repentis*. The comparison of

sequence, structure and the active site of CaDI presented the similarity to EstA, butyrylcholinesterase and acetylcholinesterase. EstA displays no detectable lipase activity on trioctanoin emulsions or olive oil nor esterase activity on cinnamoyl, feruolyl or pNP substrates. It hydrolyzes vinyl esters and triacylglycerols efficiently. EstA has a marked preference for short chain acyl moieties with orders acetate>propionate>>butyrate and triacetin>>tripropionin>tributyrin (Bourne *et al*., 2004). Therefore, the activity of CaDI might be butyrylcholinesterase and acetylcholinesterase.

The volume of the active site pocket of CaDI was 5354.5Å^3. The volume of the active site pocket of EstA was calculated to be 1555.1Å^3 using the CASTP algorithm (Dundas *et al*., 2006). The volumes of the possible substrate-binding sites of CaA1 and CaC4 were 1857.3 and 28.6Å^3, respectively. The pocket volume of the active site pocket of CaDI was larger than that of EstA, CaA1 or CaC4 and suggested its activity on complicated macromolecular carboxylesters.

Proteinase

Cloning, phylogeny and expression analyses of PrDI: The subtilase-like serine proteinase was thought to be the most important hydrolytic enzyme in the parasitism of nematophagous fungi. A serine proteinase gene, PrDI of *D. cionopaga*, was cloned and characterized. PrDI was 1359 bp in length and encoded 452 amino acids. The genomic DNA of the PrDI (GenBank accession No. JF699729.1) of *D. cionopaga* did not contain any introns. Southern blot showed that one copy of PrDI was located in the genome of *D. cionopaga* (Fig. 3). PrDI was homologous with other serine proteinase genes from nematophagous fungi. Yang *et al*. (2007a) indicated that the serine proteases from the nematophagous and entomopathogenic fungi had one common ancestor and evolved separately from those of nematode-trapping fungi. Phylogenetic analyses here showed that PrDI was clustered into a clade of proteinase genes of trapping fungi (Fig. 10). The differential expression of PrDI was analyzed by quantitative RT-PCR. PrDI was expressed at a 4-fold higher rate than in the control, indicating the role of the gene in the parasitism. The differential transcription of PrDI decreased quickly with inoculation time (Fig. 5), suggesting that the gene played a role in early infection of the nematodes.

Physicochemical properties: Signal P showed that PrDI possessed a signal peptide and the probable cleavage site was between amino acids 21 and 22. The theoretical pI/MW of the precursor would be 7.95/47.9 kDa, that of the propeptide which was indicated by comparison with the mature proteinases of other fungi, 7.37/45.7 kDa and the mature protein 5.40/33.3 kDa. The molecular weight of PrDI was less than PII and Aoz1 from *A. oligospora* with the molecular weight of 35 and 38 kDa, but similar to the 33.5 kDa PLP of *P. lilacinus* (Tunlid *et al*., 1994; Bonants *et al*., 1995; Minglian *et al*., 2004). The pI of PrDI was smaller than that of PLP by over 10.2 while higher than that of PII (4.6) and Aoz1 (4.9) (Tunlid *et al*., 1994; Bonants *et al*., 1995; Minglian *et al*., 2004). The active sites of Ser, Asp and His, determined by alignment with other serine proteinases showed that PrDI belongs to the proteinase K family. PrDI contains the possible N-glycosylation sites, protein kinase C phosphorylation sites, casein kinase II phosphorylation sites, N-myristoylation sites and the crystallins beta and gamma 'Greek key' motif signature (Table S1). However, these sites do not exist in the protein based on the developed model by PDBsite scan. The positively and negatively charged residue numbers of PrDI was 21/26. The extinction coefficient (280 nm) of the protein was 50880 M^{-1} cm^{-1}.

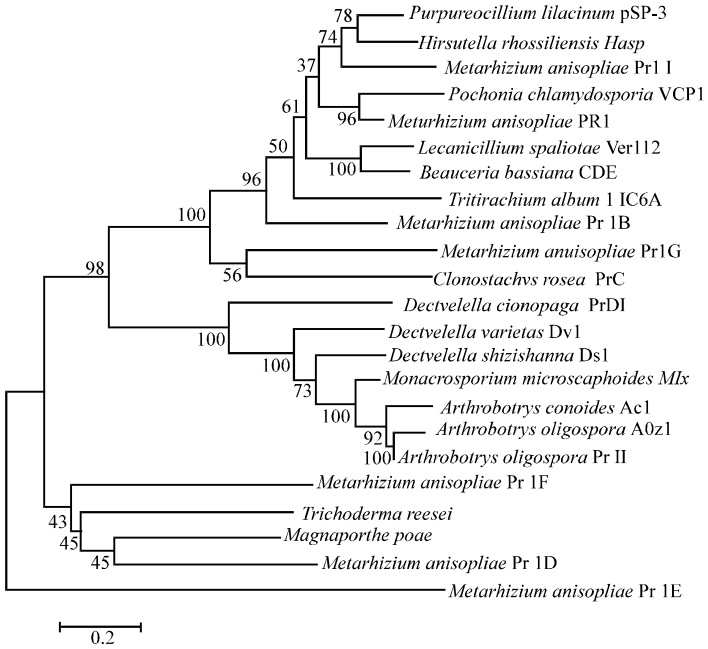

Fig. 10: Phylogeny of PrDI with the program MEGA. *Lecanicillium psalliotae* Ver112, AAU01968; *P. chlamydosporia* VCP1, CAD20584; *Meturhizium anisopliae* PR1, AAA33417; *Tritirachium album*, 1IC6A; *D. cionopaga*, AEP02888.1; *Purpureocillium lilacinum* pSP-3, ABO32256; *C. rosea* PrC, ACS66684.1; *Dactylella shizishanna* Ds1, ABL74286.1; *A. conoides* Ac1, AY859782; *D. varietas* Dv1, DQ531603; *H. rhossiliensis* Hasp, DQ422145; *M. anisopliae* Pr1I, CAB64346; *Beauveria bassiana* CDE, AAD29255; *M. anisopliae* Pr1B, CAB95012; *M. anisopliae* Pr1G, CAB63912; *M. poae*, AAD26255; *A. oligospora* Aoz1, AF516146; *A. oligospora* P¢ò, X94121; *M. microscaphoides* Mlx, AY841167; *M. anisopliae* Pr1E, CAB63915; *M. anisopliae* Pr1F, CAB63916; *M. anisopliae* Pr1D, CAB89873; *T. reesei*, EGR50513

Model description and structural comparison: The structure of PrD was modeled with 1IC6 as the template, to compare functional differences from serine proteinases. The achieved structure with 1IC6 as the template consisted of 6 α helixes and 17 β strands (Fig. 11a). A Ramachandran plot indicated that 5 amino acids of PrDI_1 were in the disallowed region. The total solvent-accessible surface area of the structure was 13158.1 Å^2. The radius of gyration was 1.77 Å. The predicted structure of PrDI indicated that the enzyme was an α/β protein, like other members of the proteinase K family. The structures VCP1 and Ver112 of nematophagous fungi and the PR1 of entomophagous fungi were constructed by homology modeling with 1IC6 as the template (Liu *et al.*, 2007). The structure of PrDI was compared with that of 1IC6 (Fig. 11b), with a RMSD of 0.32, Ver112 with 0.39, PR1 with 0.41 and VCP1 with 0.46 (Table 3) (Liu *et al.*, 2007). The RMSD calculation result indicated that PrDI was more similar to 1IC6 than to Ver112, PR1 and VCP1 (Table 3) and that it did not show the parasitic relationship. The putative Ca^{2+} binding sites of PrDI_1 were superimposed with those of PR1, VCP1 and Ver112 (Fig. 12). The oxyanion hole residue of PrDI was predicted to be Asn180. The alignment of PrDI, 1IC6, PR1, VCP1 and Ver112 was shown in Fig. 13. According to the sequence alignment, the Polar Surface Loop (PSL) of PrDI, found in the structure of other serine proteinases, consisted of the amino acids

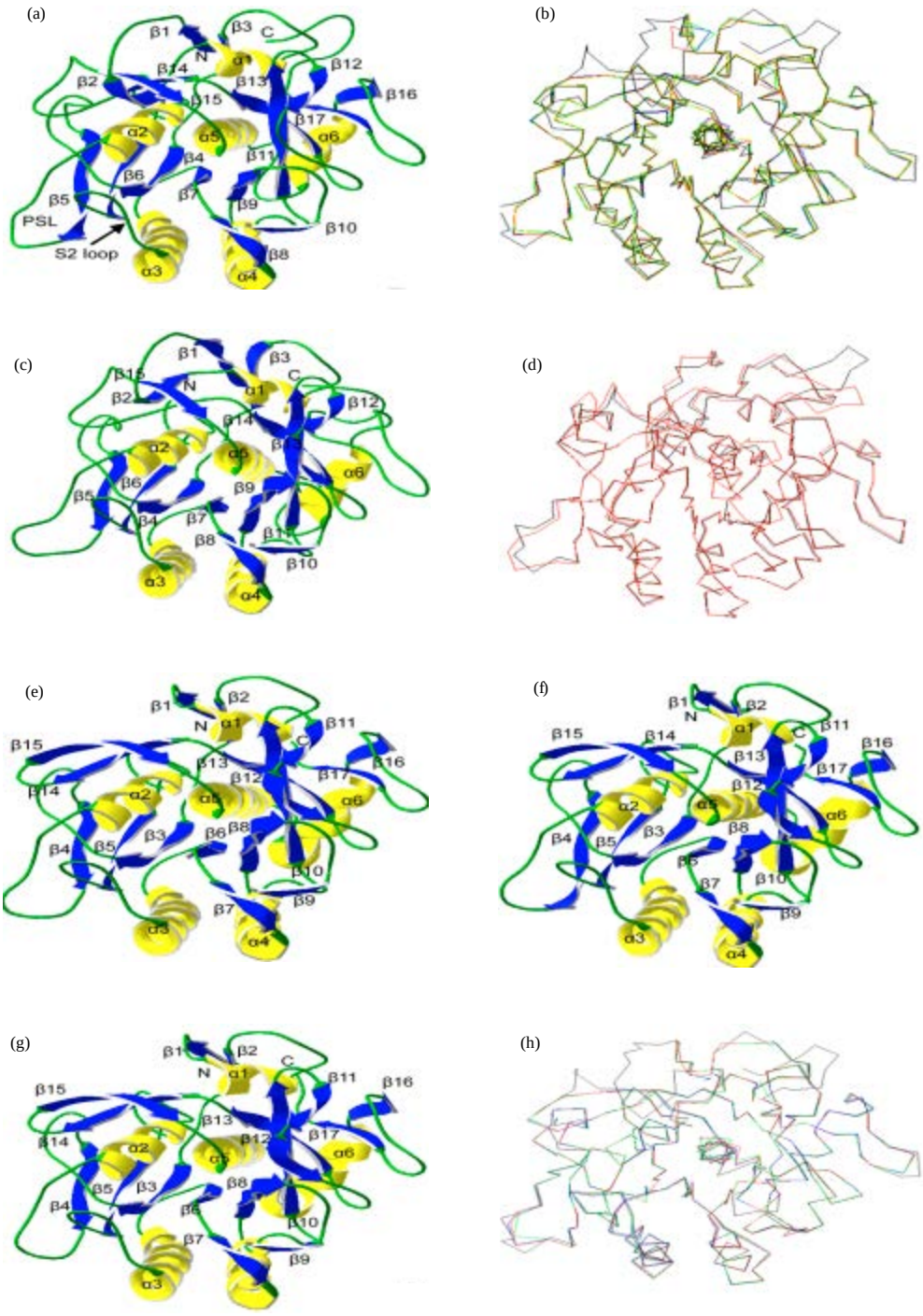

Fig. 11(a-h): Structural comparison of PrDI and homologs. (a) PrD_1, (b) Superimposition of PrDI_1 (black), PR1 (blue), VCP1 (red) and Ver112 (green), (c) PrD_ I, (d) Superimposition of PrDI_i (black) and 1P7W (red), (e) PR1_I, (f) VCP1_I, (g) Ver112_i. (h) Superimposition of PrDI_i (black), PR1_i (blue), VCP1_i (red) and Ver112_i (green)

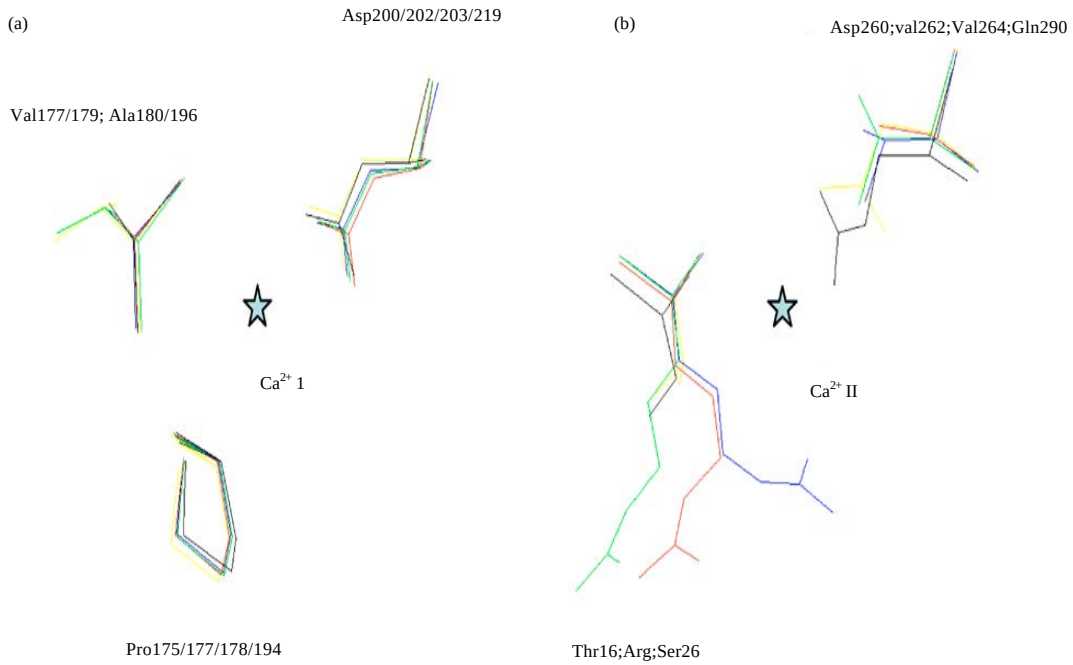

Fig. 12(a-b): Superimposition of the calcium binding sites of PrDI, 1IC6, PR1, VCP1 and Ver112. (a) Ca1 site, (b) Ca2 site. Ca2 site existing only in 1IC6. The aligned amino acids of PrDI, PR1, VCP1 and Ver112 are displayed to show the differences. PrDI, black; 1IC6, yellow; PR1, blue; Ver112, green; VCP1, red

76-NAIDTDNFD-84. However, this PSL lacked the first amino acid in β5. The chief sequence differences between PrDI and the other four proteinases were the several insertions in the N- and C-termini. The insertion YQTGKT of PrDI which constituted part of β5 and were mainly polar amino acids, was located at the N-termini of PSL. The insertion KNNNIAT in the α5-α6 loop was the second polar loop of PrDI. Structural superimposition indicated that the differences between PrDI_1 and 1IC6, VCP1, PR1 and Ver112, were in the N- and C- termini, PSL and the β16-β17 loop.

The model of PrDI that presented the active conformation was developed using proteinase K (PDB: 1P7WA) crystallized with an inhibitor of N-Pro-Ala-Pro-Phe-Ala-Ser-Ala-C. 15 β strands and 6 α helixes constituted PrDI_i (Fig. 11c). β4, β5, β6, β7, β9 and β11 were present in the saddle form. α2 and α5 were almost parallel to α3 and α4 while α1 and α6 were on a lane, perpendicular to α2, α3, α4 and α5. A Ramachandran plot indicated that 8 amino acids of PrDI_i were in the disallowed region (Fig. 6). The RMSD between PrDI_i and the template was 0.49 Å. A Ramachandran plot and RMSD analysis indicated that the PrDI structure constructed using 1P7W as the template could be applied to the docking of the inhibitor. The secondary structures of PrDI_i were highly similar to those of PrDI_1, but the model lacked the strands at the position of β14 (Val227-Ser229), β16 (Ile287-Asn289) and β17 (Ile301-Asn303) of PrDI_1. The Phe211-Asn213 of PrDI_i formed an additional strand, β13. In addition, the β1 of PrDI_1 was shorter. Structural comparison showed that the RMSD between PrDI_i and PrDI_1 was 0.6Å, the Z-score between PrDI_1 and 1IC6 and 2PWA was 54.3 and the Z-score between PrDI_1 and 1P7W and 1P7V was 54.0 while the Z-score

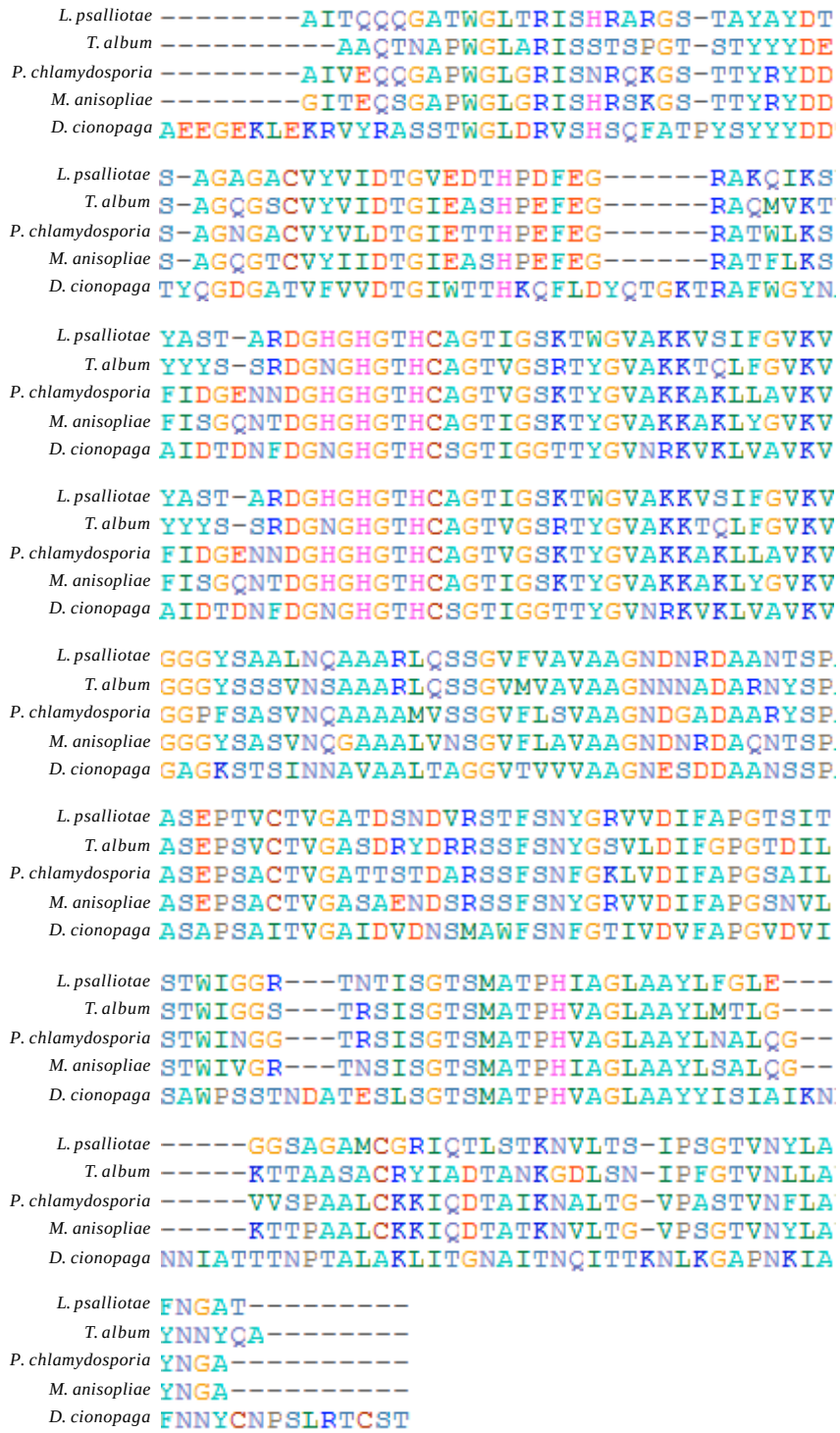

Fig. 13: Alignment of PrDI, PR1, VCP1, Ver112, and 1IC6. *L. psalliotae* Ver112, AAU01968; *T. album* 1IC6; *P. chlamydosporia* VCP1, CAD20584; *M. anisopliae* PR1, AAA33417; *D. cionopaga* PrDI, AEP02888.1

between PrDI_i and 1P7W and 1P7V was 54.3 and the Z-scores between PrDI_i and 1IC6 and other proteinase K structures (2PWA, 3GT4, 3Q40, 3GT3, 3L1K, 3Q5G and 2V8B) were 53.8. The crystallized structure of Ver112 (3F7M) and PL646 (3F7O), the cuticle-degrading enzyme, were reported. The Z-score between PrDI_1 and the crystallized structure of Ver112 (3F7M) and PL646 (3F7O) was 50.6 and 50.2, respectively and the Z-score between PrDI_i and 3F7M and 3F7O was 50.4 and 49.9, respectively.

We also modeled the structure of PR1 (GenBank: S22387), VCP1 (Gen-Bank: AJ427460) and Ver112 (GenBank: AY692148) with the template of 1P7W; to compare the differences in the substrate binding sites of PrDI and the proteinases. The structures were named PR1_i (Fig. 11e), VCP1_i (Fig. 11f) and Ver112_i (Fig. 11g), respectively. A Ramachandran plot indicated that 1 amino acid of Ver112_i, 0 of PR1_i and 0 of VCP1_i, were in the disallowed region. The RMSDs between PR1_i, VCP1_i and Ver112_i and the templates were 2.53, 5.07 and 1.60 Å, respectively. PR1_i, VCP1_i and Ver112_i all contained 6 α helixes and 17 β strands.

The variety of PrDI_i from 1P7W, VCP1_i, PR1_i and Ver112_i was in the β4-β5, β14-β15 and the α5-α6 loops (Fig. 11d and h) which were corresponding to the sequences of PrDI_1. PrDI did not contain the disulfide bridges aligned with the other four proteinases: 1IC6, PR1, VCP1 and Ver112 which contributes to the thermal stability of enzyme structure and increases the flexibility of S1 and S4 pockets located at the substrate-binding site (Liang *et al*., 2011). The possible hydrogen bonds and salt bridges of PrDI_1 were calculated to be 201 and 126 while the possible hydrogen bonds and salt bridges of PrDI_i were 201 and 105.

Substrate-binding sites and comparison: The charge analysis indicated that the active site Asp51 of PrDI was negative. The active site Asp51 was in the secondary structure the β4 of PrDI_1 and PrDI_i. The active sites formed the salt bridge Asp51:His88 that linked β4 and α2 which were also connected by Asp51: His91. Hydrogen bonds stabilize the structure of proteins. The possible hydrogen bond network Thr52/Leu115-[Asp51, Val112, Val49, Val110, Val47, Lys108, Ala45, Gly42] connected β4 with the β3-β4, β4-β5 loops, β6 with the β6-α3 loop. In addition, Asp51 might form the salt bridge with Lys113 which connected β4 and β6. His88 was in the secondary structure α2 of PrDI_1 and PrDI_i. His88 may form salt bridges with Asp84, Asp203, or Asp226 and connect α2 to the β5-α2 loop (PSL), β11 or the β14-β15 loop. The amino acid also might involve in the possible hydrogen bond network Ile96-[Cys92, His88, Asp51/Asn86] which stabilized α2. The active site Ser246 was in the secondary structure α5 of PrDI_1 and PrDI_i. Ser246-[His88, Asp51/Asn86] connected α5 and α2.

The S2-loop of PrD_1 was identified based on structural analysis and sequence alignment with VCP1, Ver112 and PR1. Two hydrogen bonds, Leu115-Asp51 and Asp118/Gly119-[Asp116, Ser120], were found to exist in the S2-loop of PrD_1. A conservative hydrogen bond network corresponding to Asp97-[Gly100, Ser101] did not appear (Liu *et al.,* 2007). In contrast to PR1, VCP1, Ver112 and 1IC6, no salt bridge was within the region of 151-SLGA-154, 158-AAGNESDDAANS-170, or 222-GTSM-225, corresponding to the S1 pocket while more hydrogen bonds (17) were within these regions. Two possible salt bridges, Arg13/Lys113:Asp116 and Arg13:Asp118, stabilized the S2-loop and linked the loop to β2 of PrD_1.

The inhibitor N-Pro-Ala-Pro-Phe-Ala-Ser-Ala-C was docked into the active sites of PrDI_i, Ver112_i, PR1_i and VCP1_i. The accomplishment of docking to the proteinases also indicated that the structures were rational. The interaction conformations were selected according to the docked energy and the cluster analysis. The electrostatic potentials and conformations of the amino acids interacting with the inhibitor are presented in Fig. 14. In the interaction of PrDI_i and the inhibitor

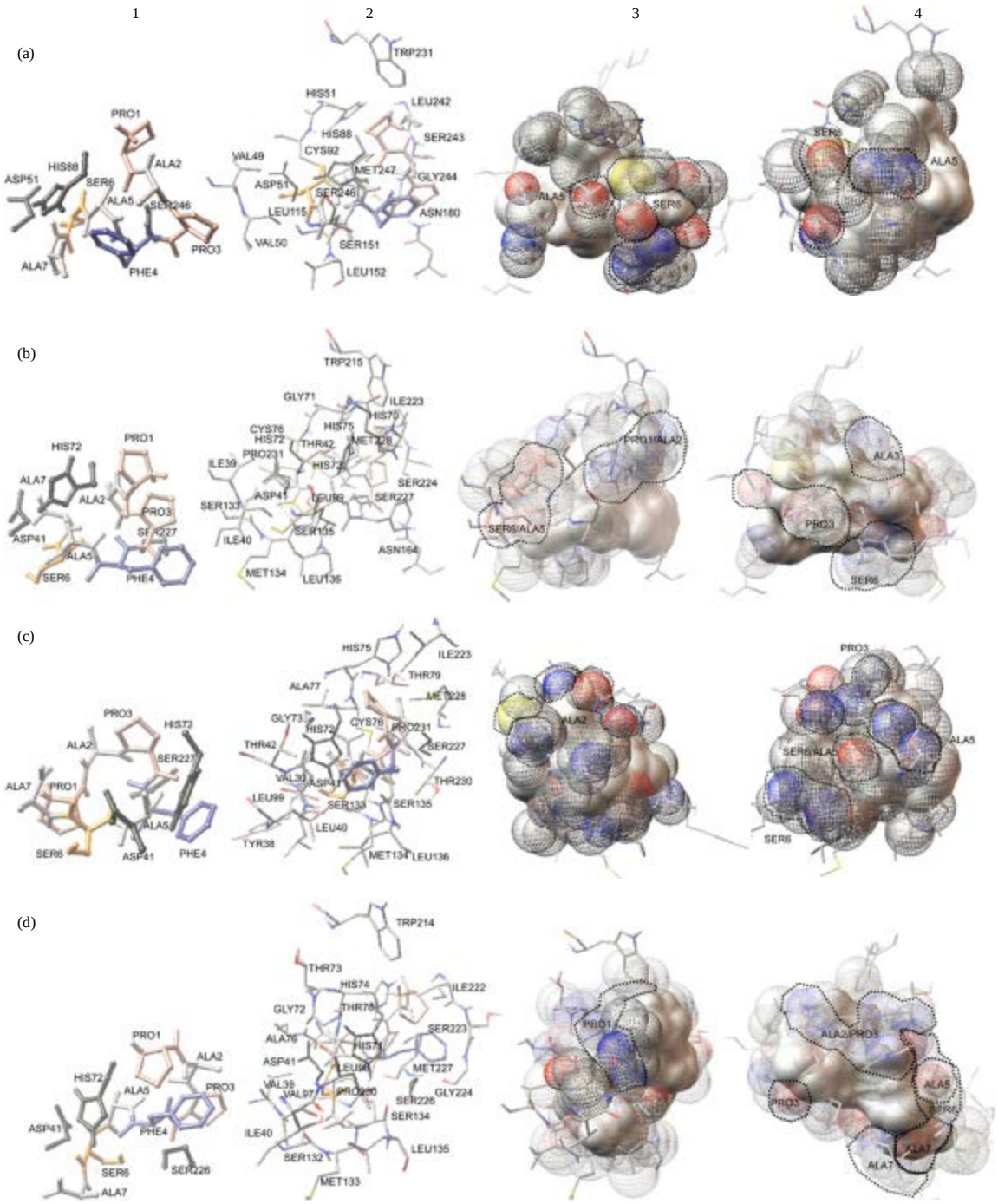

Fig. 14(a-d): Active site comparison of PrDI and homologs, (a) PrD_I, (b) PR1_i, (c) VCP1_I and (d) Ver112_i. Column 1, active sites coordinating with the inhibitor; Column 2, amino acids interacting with the heptapeptide inhibitor of Pro-Ala-Pro-Phe-Ala-Ser-Ala; Column 3, the electrostatic potential of the amino acids in contact with one side of the inhibitor; Column 4, the electrostatic potential of the amino acids in contact with the other side of the inhibitor

Table 5: Amino acids involved in the interaction of the proteinases and the inhibitor

Protein	Inhibitor/active sites		Amino acids interacted with inhibitor/active sites
PrDI_i		Pro1	Trp231, Leu242, His91, Met247 and His88
		Ala2	Leu242 and Met247
		Pro3	Ser243, Asn180, Gly244, Ser246 and Met247
	Inhibitor	Phe4	Ser151, Ser246 and His88
		Ala5	Ser246, His88 and Ser151
		Ser6	Ser151, Asp51, Cys92, His88, Val49 and Val50
		Ala7	Val50, Leu152, Ser151, Leu115, Asp51 and His88
		Asp51	Ser6 and Ala7
	Active sites	His88	Pro1, Phe4, Ala5, Ser6 and Ala7
		Ser246	Pro3, Phe4 and Ala5
Ver112_i		Pro1	Trp214, His74 and Ile222
		Ala2	His74, Ile222 and Met227
		Pro3	His74, Met227, Ile222, Thr78, Pro230, Cys75 and Ser226
	Inhibitor	Phe4	Ile222, Ser223, Gly224, Met227, Cys75, His71, Ser226 and Ser134
		Ala5	Thr73, Cys75, Ala76, Gly72, Val39, Asp41, His71 and Pro230
		Ser6	Val39, Asp41, Ser134, His71, Cys75 and Gly72
		Ala7	Ile40, Leu98, Leu135, Ser134, Ser132, Met133, Val97 and Asp41
		Asp41	Ala5, Ser6 and Ala7
	Active sites	His72	Phe4, Ala5 and Ser6
		Ser226	Pro3 and Phe4
Pr1_i		Pro1	Trp215, Ile223, Gly71, His70 and His72
		Ala2	His75, Met228 and Cys76
		Pro3	Asn164, Ser224, Ser227 and Met228
	Inhibitor	Phe4	Asn164, Ser135, Leu136 and Ser227
		Ala5	Asp41, Thr42 and Leu99
		Ser6	Ile40, Asp41, Met134, Ser135 and Leu136
		Ala7	Ile39, Asp41, Cys76 and Pro231
		Asp41	Ala5 and Ser6
	Active sites	His72	Pro1
		Ser227	Pro3 and Phe4
VCP1_i		Pro1	Thr230, Ser133, Ser135, Met134 and Pro231
		Ala2	Met228, Ser227, Thr79, Pro231, Cys76 and His75
		Pro3	His75, Ile223, Met228, Cys76 and Ser227
	Inhibitor	Phe4	Thr42, Leu99, Ser135, Leu136, Ser227 and His72
		Ala5	His72, Asp41, Gly73, Cys76, Thr42 and Ser135
		Ser6	Tyr38, Ser135, Ser133, Leu40, Val39, Asp41 and Cys76
		Ala7	Ser133, Pro231, Val39, Cys76, Ala77 and Gly73
		Asp41	Ala5 and Ser6
	Active sites	His72	Phe4 and Ala5
		Ser227	Ala2, Pro3 and Phe4

N-Pro-Ala-Pro-Phe-Ala-Ser-Ala-C, no hydrogen bond was found to exist between the protein and the inhibitor. The structural analysis of PrDI indicated that the active sites of the enzyme and the binding sites of the inhibitor formed one pocket. The binding energy was 49.42 kcal mol^{-1}. Val49, Val50, Asp51, His88, His91, Cys92, Leu115, Ser151, Leu152, Asn180, Trp231, Leu242, Ser243, Gly244, Ser246 and Met247; involved in the interaction of the inhibitor and PrDI (Table 5).

Table S1: Putative post-translational modification sites of CaDI and PrDI

Putative post-translational modification sites	Location
CaDI	
Asn-linked glycosylation sites	31N-34S, 100N-103A, 531N-534G
Amidation site	63V-K66, 346N-349R
cAMP-and cGMP-dependent protein kinase phosphorylation site	65K-68T
Casein kinase II phosphorylation site	34S-37D, 157T-160E, 181S-184D, 385S-388D, 409T-412D, 515T-518E
N-myristoylation site	41G-46V, 116G-121L, 142G-147F, 189G-194G, 213G-218K, 223G-228G, 239G-244T, 280G-285S, 316G-321W, 421G-426G, 482G-487L, 500G-505S, 534G-539T
Protein kinase C phosphorylation site	51S-53R, 68T-70R, 129S-131K, 157T-159K, 342S-344K, 372S-374K, 400S-402K, 452S-454K, 515T-517R
Tyrosine kinase phosphorylation site	453K-459Y
PrDI	
N-glycosylation sites	57T-59K, 66T-68K, 292T-294K, 312S-314R
casein kinase II phosphorylation sites	125S-128D, 216T-219D, 234S-237D
N-myristoylation sites	53G-58H, 85G-90T, 94G-99T, 121G-126V, 153G-158T, 244G-249T, 284G-289N

The subsites of the substrate-binding regions of proteins can be assigned by the interaction characterization of the inhibitor and the enzyme (Schechter and Berger, 1967). At the contact of N-Pro-Ala-Pro-Phe-Ala-Ser-Ala-C and PrDI_i, Pro1 of the inhibitor interacted closely with Trp231, Leu242, His91, Met247 and His88 of the protein. Ala2 interacted with Leu242 and Met247. The interaction of Pro3 and PrDI_i was mediated by Ser243, Asn180, Gly244, Ser246 and Met247. Subsite S1 that interacted with Phe4 consisted of Ser151, Ser246 and His88 while subsite S1' that interacted with Ala5 consisted of Ser246, His88 and Ser151. Ser151, Asp51, Cys92, His88, Val49 and Val50 constituted the subsite S2' interaction with Ser6. The amino acids that interacted with Ala7 included Val50, Leu152, Ser151, Leu115, Asp51 and His88.

The active site Asp51 of PrDI_i interacted with Ser6 and Ala7 of the inhibitor, His88 interacted with Pro1, Phe4, Ala5, Ser6 and Ala7 and Ser246 interacted with Pro3, Phe4 and Ala5. The conformation of the active sites of PrDI_i changed upon binding the inhibitor. CA of His88 was towards the α-1-pyrrolidinyl of Pro1 while ND1 towards the aromatic ring of Phe4 (Fig. 14). The alkaline protease PrDI can form the correct protonation state which is guaranteed by a high pH value (Liang *et al*., 2010). The amino acids interacting with Ala5 and Ser6 showed the differing electrostatic potential on both sides (Fig. 14). The amino acids interacting closely with Pro1 and Ala2 were hydrophilic while those with Ala7 were mainly hydrophobic which was presented by ProScale.

In the interaction of Ver112_i and the inhibitor, the amino acids of Val39, Ile40, Asp41, His71, Gly72, Thr73, His74, Cys75, Ala76, Thr78, Val97, Leu98, Ser132, Met133, Ser134, Leu135, Trp214, Ile222, Ser223, Gly224, Ser226, Met227 and Pro230 were involved (Table 5). The binding energy was 2750 kcal mol^{-1}. Hydrogen bonds of Cys75:Phe4 and Ser134:Pro3 existed between the protein and the inhibitor. Consistent with that of PrDI_i, the conformation of the active sites of Ver112_i, PR1_i and VCP1_i was transformed when the proteins were docked with the inhibitor. In the docking complex of Ver112_i and the inhibitor, the β-imidazole of His72 of Ver112_i was almost parallel to the α-1-pyrrolidinyl of Pro1 of the inhibitor and the aromatic ring of Phe4. CA and CB of Ser226 were towards the aromatic ring of Phe4 (Fig. 14). The amino acids interacting with Pro1 had positive electrostatic potential while those with Ser6 had mainly negative

electrostatic potential (Fig. 14). The amino acids interacting with Ala2 on one side had positive electrostatic potential. The amino acids interacting with Pro3, Ala5 and Ala7 had positive electrostatic potential on one side and negative on the other.

In the interaction of PR1_i and the inhibitor, a hydrogen bond existed between Ser135 of the protein and Ala5 of the inhibitor. The amino acids involved in the interaction between the two molecules were Ile39, Ile40, Asp41, Thr42, His70, Gly71, His72, His75, Cys76, Leu99, Ser133, Met134, Ser135, Leu136, Asn164, Trp215, Ile223, Ser224, Ser227, Met228 and Pro231 (Table 5). The binding energy of PR1_i and the inhibitor was 220.34 kcal mol^{-1}. CD2 of His72, CA and CB of Ser227, were, respectively situated towards the α-1-pyrrolidinyl of Pro1 and Pro3. The amino acids interacting with Pro1 and Ala2 showed a positive electrostatic potential patch and the amino acids interacting with Pro3 showed a negative electrostatic potential patch. In contrast, the amino acids interacting with Ser6 and Ala5 showed a negative electrostatic potential patch on one side and a positive electrostatic potential patch on the other.

In the interaction of VCP1_i and the inhibitor, the amino acids of Tyr38, Val39, Leu40, Asp41, Thr42, His72, Gly73, His75, Cys76, Ala77, Thr79, Leu99, Ser133, Met134, Ser135, Leu136, Ile223, Ser227, Met228, Thr230 and Pro231 were implicated (Table 5). The binding energy was 1330 kcal mol^{-1}. The β-imidazole of His72 was almost parallel to the aromatic ring of Phe4. The amino acid interacting with Ala2 presented two negative patches and two positive patches. The amino acid interacting with Pro3 was mainly positive. The amino acid interacting with Ala5 and Ser6 was negative on one side and positive on the other.

The inhibitor-binding pockets of PrDI_i, PR1_i, VCP1_i and Ver112_i are shown in Fig. 15. The active site alignment of these proteinases was showed in Fig. 16. The differences of the binding modes of substrates and key amino acid residues in active site regions can facilitate virulence improvements of serine proteinases for fungi against nematodes or insects. From the figure, V50 was the specific amino acids of PrDI in the active site regions included which interacted with the inhibitor. PrDI presented less amino acids needed by the interaction with the inhibitor except N180. The conserve amino acid Asn also involved in the interaction between PR1 and the inhibitor.

Liu *et al.* (2007) proposed that differences in the electrostatic surface potential, the hydrophobicity and size of the S4 substrate-binding pocket and the number and distribution of hydrogen bonds and salt bridges within regions that were part of or in close proximity to the S2-loop, led to the functional variety of the serine proteinases of nematophagous fungi. The electrostatic potential around the active sites of PrD_1 was negative like Ver112 and PL646 (Liang *et al.*, 2010). The surface of PrDI was mainly positively charged at neutral and alkaline pH values like P32, Ver112 and PL646, allowing the protein to bind to predominant negatively charged residues in the nematode eggshell and the nematode female (Larriba *et al.*, 2012; Liang *et al.*, 2010). However, the putative S4 pocket of PrDI_1 did not present two subsites (Fig. 17). The residues that consisted of the pocket were hydrophobic or hydrophilic. These characterizations were not consistent with the hydrophobic properties of the S4 pocket or the electrostatic potential in the active site pockets of VCP1 and PR1 (Liu *et al.*, 2007). There was not the conserve Tyr in the S4 pocket like Ver112, the side chain of which is a large phenoxyl ring that is located in the external region of the S4 pocket and acts as a lid to this pocket, restricting to a certain extent the binding of the peptide substrate with large P4 side chains (Liang *et al.*, 2010).

In comparison with the results of Liu *et al.* (2007) we found that the inhibitor-binding sites of these proteinases were different which was consistent with the results of Larriba *et al.* (2012) and Liang *et al.* (2010). The comparison of the substrate-binding sites of PrDI_i, Ver112_i and VCP1_i

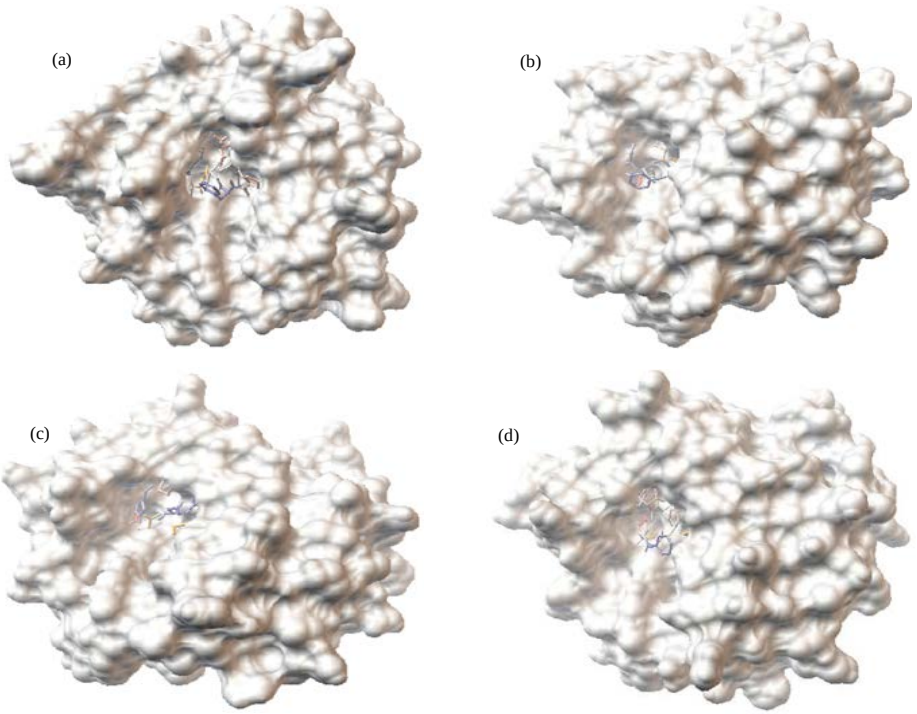

Fig. 15(a-d): Binding pockets of the heptapeptide inhibitor of Pro-Ala-Pro-Phe-Ala-Ser-Ala of PrDI_i, VCP1_i, Ver112_i and PR1_i. (a) PrDI_I, (b) VCP1_I, (c) Ver112_I and (d) PR1_i

```
Ver112 CVYVIDTGVEDTHPDFE-----GRAKQIKSYAS--T-ARDGHGHGTHCAGTIGSKTWGVAKKVSIFGVKVLDD 100
VCP1   CVYVLDTGIETTHPEFE-----GRATWLKSFID—GENNDGHGHGTHCAGTVGSKTYGVAKKAKLLAVKVLDN 101
PR1    CVYIIDTGIEASHPEFE-----GRATFLKSFIS—GQNTDGHGHGTHCAGTIGSKTYGVAKKAKLYGVKVLDN 101
PrDI   TVFVVDTGIWTTHKQFLDYQTGKTRAFWGYNAIDTDNFDGNGHGTHCSGTIGGTTYGVNRKVKLVAVKVLDA 117

Ver112 SGSGSLSNIVAGMDFVASDRQSRNCPRRTVASMSLGGGYSAALNQAAARLQSSGVFVAVAAGNDNRDAANTS 172
VCP1   GGSGSYAGVIAGMEFVSQDYKTRGCPNGAIASMSLGGPFSASVNQAAAAMVSSGVFLSVAAGNDGADAARYS 173
PR1    QGSGSYSGIISGMDYVAQDSKTRGCPNGAIASMSLGGGYSASVNQGAAALVNSGVFLAVAAGNDNRDAQNTS 173
PrDI   DGSGSLESVIDGMQWVLKNVTSLGLQNKAVVSMSLGAGKSTSINNAVAALTAGGVTVVVAAGNESDDAANSS 189

Ver112 PASEPTVCTVGATDSNDVRSTFSNYGRVVDIFAPGTSITSTWIGGR---TNTISGTSMATPHIAG 234
VCP1   PASEPSACTVGATTSTDARSSFSNFGKLVDIFAPGSAILSTWINGG---TRSISGTSMATPHVAG 235
PR1    PASEPSACTVGASAENDSRSSFSNYGRVVDIFAPGSNVLSTWIVGR---TNSISGTSMATPHIAG 235
PrDI   PASAPSAITVGAIDVDNSMAWFSNFGTIVDVFAPGVDVISAWPSSTNDATESLSGTSMATPHVAG 254
```

Fig. 16: Alignment of amino acids interacting with the inhibitor of PrDI, PR1, VCP1 and Ver112

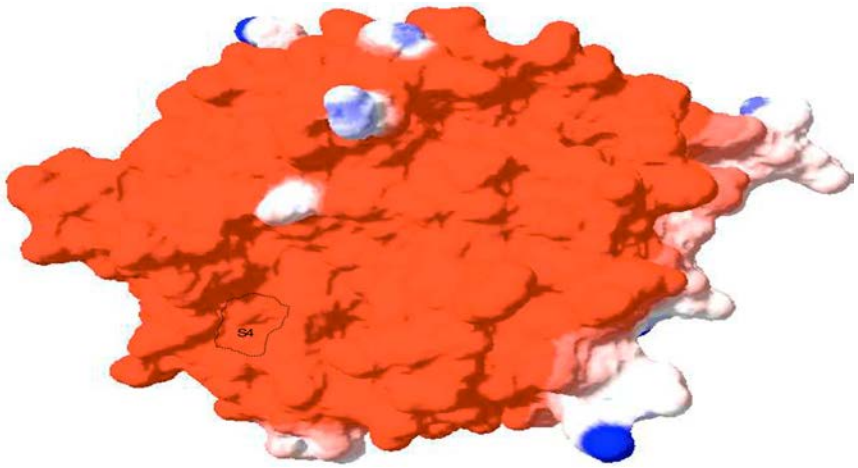

Fig. 17: Binding subsite, S4, of the inhibitor Methoxysuccinyl-Ala-Ala-Pro-Ala-chloromethyl ketone of PrDI_1

showed the conservation of the catalytic region among the serine proteinases of nematophagous fungi (Fig. 16). However, the specific amino acids in contact with the residues of the inhibitor were still various and the electrostatic potential difference of the binding pocket and the conformations of the active sites all showed the diversity of enzymes (Fig. 14).

CONCLUSION

The possible infection-related enzymes PrDI and CaDI of *D. cionopaga* were cloned and characterized using the analysis of phylogeny, differential transcription and structural comparison. Quantitative RT-PCR confirmed that these genes were transcripted more highly during the infection of nematodes. The polar, hydrophilic and large active site region suggested that CaDI probably catalyzed the long-chain carboxylester in comparison to EstA. The putative protein of PrDI was clustered into a clade with the proteinases of nematophagous, trapping fungi. Structural studies of protease-inhibitor complexes can provide information on the binding modes of substrates and key amino acid residues associated with catalytic domains Liang *et al.* (2010). To present its structural difference from the known models of serine proteinases of nematophagous and entomophagous fungi, PrDI was modeled with two templates and the docking of the heptapeptide inhibitor was performed. The computation results of the conformation, electrostatic potential, hydrophobicity and charge of the amino acids in the active site pockets indicated that the hydrolyzing subsites of the cuticle-degrading proteinases were conservative and specific. The results of the transcription analysis and the sequence and structural similarity analysis suggested that the putative proteins encoded by PrDI and CaDI are involved in the parasitism of the nematophagous fungus *D. cionopaga*.

ACKNOWLEDGMENT

This study was sponsored by the Postdoctoral Science Foundation of China (20070410155).

REFERENCES

Adinarayana, K.P.S. and R.K. Devi, 2011. Protein-Ligand interaction studies on 2, 4, 6-trisubstituted triazine derivatives as anti-malarial DHFR agents using AutoDock. Bioinformation, 6: 74-77.

Ahman, J., B. Ek, L. Rask and A. Tunlid, 1996. Sequence analysis and regulation of a gene encoding a cuticle-degrading serine protease from the nematophagous fungus *Arthrobotrys oligospora*. Microbiology, 142: 1605-1616.

Ahren, D., M. Tholander, C. Fekete, B. Rajashekar, E. Friman, T. Johansson and A. Tunlid, 2005. Comparison of gene expression in trap cells and vegetative hyphae of the nematophagous fungus *Monacrosporium haptotylum*. Microbiology, 151: 789-803.

Angkawidjaja, C., Y. Koga, K. Takano and S. Kanaya, 2012. Structure and stability of a thermostable carboxylesterase from the thermoacidophilic archaeon *Sulfolobus tokodaii*. FEBS J., 279: 3071-3084.

Betzel, C., S. Gourinath, P. Kumar, P. Kaur, M. Perbandt, S. Eschenburg and T.P. Singh, 2001. Structure of a serine protease proteinase K from *Tritirachium album limber* at 0.98 Å resolution. Biochemistry, 40: 3080-3088.

Bonants, P.J., P.F. Fitters, H. Thijs, E. den Belder, C. Waalwijk and J.W.D. Henfling, 1995. A basic serine protease from *Paecilomyces lilacinus* with biological activity against Meloidogyne hapla eggs. Microbiology, 141: 775-784.

Bourne, Y., A.A. Hasper, H. Chahinian, M. Juin, L.H. de Graaff and P. Marchot, 2004. *Aspergillus niger* protein EstA defines a new class of fungal esterases within the α/β hydrolase fold superfamily of proteins. Structure, 12: 677-687.

Brendel, V., P. Bucher, I.R. Nourbakhsh, B.E. Blaisdell and S. Karlin, 1992. Methods and algorithms for statistical analysis of protein sequences. Proc. Natl. Acad. Sci. USA., 89: 2002-2006.

Cayrol, J.C., 1983. [Biological control of meloidogyne by *Arthrobotrys irregularis*]. Revue Nematol., 6: 265-273, (In French).

Cayrol, J.C., J.P. Frankowski, A. Laniece, G. D'Hardemare and J.P. Talon, 1978. [Against nematodes in mushroom. Development of a method of biological control using a predator Hyphomycete: *Arthrobotrys robusta* strain antipolis]. PHM-Revue Horticole, 184: 23-30, (In French).

Colombo, A.L., A.S.A. Melo, R.F.C. Rosas, R. Salomao and M. Briones *et al.*, 2003. Outbreak of *Candida rugosa* candidemia: An emerging pathogen that may be refractory to amphotericin B therapy. Diagn. Microbiol. Infect. Dis., 46: 253-257.

Davide, R.G. and R.A. Zorilla, 1983. Evaluation of a fungus, *Paecilomyces lilacinus* (Thom) Samson, for the biological control of the potato cyst nematode, *Globodera rostochiensis* Woll. as compared with some inmaticides. Philippine Agric., 66: 397-404.

De Simone, G., V. Menchise, V. Alterio, L. Mandrich, M. Rossi, G. Manco and C. Pedone, 2004. The crystal structure of an EST2 mutant unveils structural insights on the H group of the carboxylesterase/lipase family. J. Mol. Biol., 343: 137-146.

Doucet-Personeni, C., P.D. Bentley, R.J. Fletcher, A. Kinkaid and G. Kryger *et al.*, 2001. A structure-based design approach to the development of novel, reversible AChE inhibitors. J. Med. Chem., 44: 3203-3215.

Dundas, J., Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz and J. Liang, 2006. CASTp: Computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. Nucleic Acid Res., 34: 116-118.

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel and A. Bairoch, 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: The Proteomics Protocols Handbook, Walker, J.M. (Ed.). 1st Edn., Humana Press, New Jersey, USA., ISBN-13: 978-1588295934, pp: 571-607.

Gill, S.C. and P.H. von Hippel, 1989. Calculation of protein extinction coefficients from amino acid sequence data. Anal. Biochem., 182: 319-326.

Grochulski, P., F. Bouthillier, R.J. Kazlauskas, A.N. Serreqi, J.D. Schrag, E. Ziomek and M. Cygler, 1994. Analogs of reaction intermediates identify a unique substrate binding site in *Candida rugosa* lipase. Biochemistry, 33: 3494-3500.

Guex, N. and M.C. Peitsch, 1997. SWISS-MODEL and the Swiss-PDB viewer: An environment for comparative protein modeling. Electrophoresis, 18: 2714-2723.

Huang, X., N. Zhao and K. Zhang, 2004. Extracellular enzymes serving as virulence factors in nematophagous fungi involved in infection of the host. Res. Microbial., 155: 811-816.

Huey, R., G.M. Morris, A.J. Olson and D.S. Goodsell, 2007. A semiempirical free energy force field with charge-based desolvation. J. Comput. Chem., 28: 1145-1152.

Kabsch, W. and C. Sander, 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. Biopolymers, 22: 2577-2637.

Kanda, S., T. Aimi, S. Kano, S. Ishihara, Y. Kitamoto and T. Morinaga, 2008. Ambient pH signaling regulates expression of the serine protease gene (*spr1*) in pine wilt nematode-trapping fungus, *Monacrosporium megalosporum*. Microbiol. Res., 163: 63-72.

Kim, M.H., B.S. Kang, S. Kim, K.J. Kim and C.H. Lee *et al*., 2008. The crystal structure of the estA protein, a virulence factor from *Streptococcus pneumoniae*. Proteins: Struct. Funct. Bioinform., 70: 578-583.

Kumar, S. and R. Nussinov, 1999. Salt bridge stability in monomeric proteins. J. Mol. Biol., 293: 1241-1255.

Kumar, S., C.J. Tsai, B. Ma and R. Nussinov, 2000. Contribution of salt bridges toward protein thermostability. J. Biomol. Struct. Dyn., 17: 79-85.

Kumar, S. and R. Nussinov, 2002. Relationship between ion pair geometries and electrostatic strengths in proteins. Biophys. J., 83: 1595-1612.

Larriba, E., J. Martin-Nieto and L.V. Lopez-Llorca, 2012. Gene cloning, molecular modeling and phylogenetics of serine protease P32 and serine carboxypeptidase SCP1 from nematophagous fungi *Pochonia rubescens* and *Pochonia chlamydosporia*. Can. J. Microbiol., 58: 815-827.

Liang, L., Z. Meng, F. Ye, J. Yang and S. Liu *et al*., 2010. The crystal structures of two cuticle-degrading proteases from nematophagous fungi and their contribution to infection against nematodes. FASEB J., 24: 1391-1400.

Liang, L., S. Liu, J. Yang, Z. Meng, L. Lei and K. Zhang, 2011. Comparison of homology models and crystal structures of cuticle-degrading proteases from nematophagous fungi: Structural basis of nematicidal activity. FASEB J., 25: 1894-1902.

Lindahl, E., B. Hess and D. Van der Spoel, 2001. GROMACS 3.0 a package for molecular simulation and trajectory analysis. J. Mol. Model, 7: 306-317.

Liu, S.Q., Z.H. Meng, J.K. Yang, Y.X. Fu and K.Q. Zhang, 2007. Characterizing structural features of cuticle-degrading proteases from fungi by molecular modeling. BMC Struct. Biol., Vol. 7. 10.1186/1472-6807-7-33

Minglian, Z., M. Minghe and Z. Keqin, 2004. Characterization of a neutral serine protease and its full-length cDNA from the nematode-trapping fungus *Arthrobotrys oligospora*. Mycologia, 96: 16-22.

Montes, M.J., I. Lopez-Brana and A. Delibes, 2004. Root enzyme activities associated with resistance to *Heterodera avenae* conferred by gene *Cre*7 in a wheat/*Aegilops triuncialis* introgression line. J. Plant Physiol., 161: 493-495.

Morris, G.M., R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell and A.J. Olson, 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J. Comput. Chem., 30: 2785-2791.

Morton, C.O., P.R. Hirsch, J.P. Peberdy and B.R. Kerry, 2003. Cloning of and genetic variation in protease VCP1 from the nematophagous fungus *Pochonia chlamydosporia*. Mycol. Res., 107: 38-46.

Nicolet, Y., O. Lockridge, P. Masson, J.C. Fontecilla-Camps and F. Nachon, 2003. Crystal structure of human butyrylcholinesterase and of its complexes with substrate and products. J. Biol. Chem., 278: 41141-41147.

Petersen, T.N., S. Brunak, G. von Heijne and H. Nielsen, 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nat. Methods, 8: 785-786.

Sarakatsannis, J.N. and Y. Duan, 2005. Statistical characterization of salt bridges in proteins. Proteins: Struct. Funct. Bioinform., 60: 732-739.

Schechter, I. and A. Berger, 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun., 27: 157-162.

Schrag, J.D. and M. Cygler, 1993. 1•8 Å refined structure of the lipase from *Geotrichum candidum*. J. Mol. Biol., 230: 575-591.

Segers, R., T.M. Butt, J.N. Keen, B.R. Kerry and J.F. Peberdy, 1995. The subtilisins of the invertebrate mycopathogens *Verticillium chlamydosporium* and *Metarhizium anisopliae* are serologically and functionally related. FEMS Microbiol. Lett., 126: 227-231.

Segers, R., T.M. Butt, B.R. Kerry, A. Beckett and J.F. Peberdy, 1996. The role of the proteinase VCP1 produced by the nematophagous *Verticillium chlamydosporium* in the infection process of nematode eggs. Mycol. Res., 100: 421-428.

St. Leger, R.J., A.K. Charnley and R.M. Cooper, 1987. Characterization of cuticle-degrading proteases produced by the entomopathogen *Metarhizium anisopliae*. Arch. Biochem. Biophys., 253: 221-232.

St. Leger, R.J., D.C. Frank, D.W. Roberts and R.C. Staples, 1992. Molecular cloning and regulatory analysis of the cuticle-degrading-protease structural gene from the entomopathogenic fungus *Metarhizium anisopliae*. Eur. J. Biochem., 204: 991-1001.

Thompson, J.D., D.G. Higgins and T.J. Gibson, 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.

Tunlid, A., S. Rosen, B. Ek and L. Rask, 1994. Purification and characterization of an extracellular serine protease from the nematode-trapping fungus *Arthrobotrys oligospora*. Microbiology, 140: 1687-1695.

Wang, B., W. Wu and X. Liu, 2007. Purification and characterization of a neutral serine protease with nematicidal activity from *Hirsutella rhossiliensis*. Mycopathologia, 163: 169-176.

Wang, B., X. Liu, W. Wu, X. Liu and S. Li, 2009. Purification, characterization and gene cloning of an alkaline serine protease from a highly virulent strain of the nematode-endoparasitic fungus *Hirsutella rhossiliensis*. Microbiol. Res., 164: 665-673.

Wang, S., W. Fang, C. Wang and R.J. St. Leger, 2011. Insertion of an esterase gene into a specific locust pathogen (*Metarhizium acridum*) enables it to infect caterpillars. PLoS Pathog., Vol. 7. 10.1371/journal.ppat.1002097

Yang, J., X. Huang, B. Tian, H. Sun, J. Duan, W. Wu and K. Zhang, 2005a. Characterization of an extracellular serine protease gene from the nematophagous fungus *Lecanicillium psalliotae*. Biotechnol. Lett., 27: 1329-1334.

Yang, J., X. Huang, B. Tian, M. Wang, Q. Niu and K. Zhang, 2005b. Isolation and characterization of a serine protease from the nematophagous fungus, *Lecanicillium psalliotae*, displaying nematicidal activity. Biotechnol. Lett., 27: 1123-1128.

Yang, Y., 2006. Taxonomy of predatory fungi of *Orbiliaceae* (Ascomycotina) and molecular evolution of trapping devices. Ph.D. Thesis, Graduate School of Chinese Academy of Sciences, China.

Yang, J., B. Tian, L. Liang and K.Q. Zhang, 2007a. Extracellular enzymes and the pathogenesis of nematophagous fungi. Applied Microbiol. Biotechnol., 75: 21-31.

Yang, J., J. Li, L. Liang, B. Tian, Y. Zhang, C. Cheng and K.Q. Zhang, 2007b. Cloning and characterization of an extracellular serine protease from the nematode-trapping fungus *Arthrobotrys conoides*. Arch. Microbiol., 188: 167-174.

Yang, J., L. Liang, Y. Zhang, J. Li and L. Zhang *et al*., 2007c. Purification and cloning of a novel serine protease from the nematode-trapping fungus *Dactylellina varietas* and its potential roles in infection against nematodes. Applied Microbiol. Biotechnol., 75: 557-565.

Yang, J., L. Wang, X. Ji, Y. Feng and X. Li *et al*., 2011. Genomic and proteomic analyses of the fungus *Arthrobotrys oligospora* provide insights into nematode-trap formation. PLoS Pathog., Vol. 7. 10.1371/journal.ppat.1002179

Yu, H., J. Duan, B. Wang and X. Jiang, 2011. The function of snodprot in the cerato-platanin family from *Dactylellina cionopaga* in nematophagous fungi. Biosci. Biotechnol. Biochem., 76: 1835-1842.

Zhao, M.L., J.S. Huang, M.H. Mo and K.Q. Zhang, 2005. A potential virulence factor involved in fungal pathogenicity: Serine-like protease activity of nematophagous fungus *Clonostachys rosea*. Fungal Divers., 19: 217-234.