ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/256501584

# Molecular cloning and homology modelling of a subtilisin-like serine protease from the marine fungus, Engyodontium album BTMFS10

**ARTICLE** *in* WORLD JOURNAL OF MICROBIOLOGY AND BIOTECHNOLOGY · JULY 2010

Impact Factor: 1.78 · DOI: 10.1007/s11274-009-0298-6 · Source: PubMed

CITATIONS
7

READS
146

**6 AUTHORS**, INCLUDING:

Jasmin C
National Institute of Oceanography

**10** PUBLICATIONS **120** CITATIONS

SEE PROFILE

Sreeja Chellappan
Indian Institute of Science

**10** PUBLICATIONS **86** CITATIONS

SEE PROFILE

Rajeev K Sukumaran
National Institute for Interdisciplinary Scie…

**85** PUBLICATIONS **1,893** CITATIONS

SEE PROFILE

Chandrasekaran Muthusamy
King Saud University

**63** PUBLICATIONS **627** CITATIONS

SEE PROFILE

Available from: Rajeev K Sukumaran
Retrieved on: 24 January 2016

World J Microbiol Biotechnol (2010) 26:1269–1279
DOI 10.1007/s11274-009-0298-6

ORIGINAL PAPER

# Molecular cloning and homology modelling of a subtilisin-like serine protease from the marine fungus, *Engyodontium album* BTMFS10

**C. Jasmin · Sreeja Chellappan · Rajeev K. Sukumaran · K. K. Elyas · Sarita G. Bhat · M. Chandrasekaran**

Received: 19 July 2009 / Accepted: 22 December 2009 / Published online: 10 January 2010
© Springer Science+Business Media B.V. 2010

**Abstract** An alkaline protease gene (*Eap*) was isolated for the first time from a marine fungus, *Engyodontium album*. *Eap* consists of an open reading frame of 1,161 bp encoding a prepropeptide consisting of 387 amino acids with a calculated molecular mass of 40.923 kDa. Homology comparison of the deduced amino acid sequence of *Eap* with other known proteins indicated that *Eap* encode an extracellular protease that belongs to the subtilase family of serine protease (Family S8). A comparative homology model of the *Engyodontium album* protease (EAP) was developed using the crystal structure of proteinase K. The model revealed that EAP has broad substrate specificity similar to Proteinase K with preference for bulky hydrophobic residues at P1 and P4. Also, EAP is suggested to have two disulfide bonds and more than two $Ca^{2+}$ binding sites in its 3D structure; both of which are assumed to contribute to the thermostable nature of the protein.

**Keywords** *Engyodontium album* · Alkaline serine protease · Subtilases · Homology modelling

**Abbreviations**

| | |
|---|---|
| *E. album* | *Engyodontium album* |
| EAP | *Engyodontium album* protease (deduced protein) |
| *Eap* | *Engyodontium album* protease (gene) |

C. Jasmin · S. Chellappan · K. K. Elyas · S. G. Bhat · M. Chandrasekaran (✉)
Microbial Technology Laboratory, Department of Biotechnology, Cochin University of Science and Technology, Cochin 682 022, India
e-mail: chansek10@gmail.com

R. K. Sukumaran
Biotechnology Division, National Institute for Interdisciplinary Science and Technology, Thiruvananthapuram 695 019, India

## Introduction

Proteases which account for about 60% of total enzyme market, is the largest class of enzymes occupying a pivotal position owing to their wide industrial and research applications such as detergent, food, pharmaceuticals, leather, diagnostics, waste management, silver recovery and molecular biology (Rao et al. 1998). Among the microbial proteases, the serine proteases are the largely exploited group consisting of 20 families (S1–S27), which are grouped into 6 clans (SA, SB, SC, SD, SF and SG) based on the structural similarities and other functional evidence. Microbial serine proteases account for approximately 40% of the total worldwide enzyme sale (Rao et al. 1998; Sharma et al. 2006). Subtilases (Subtilisin like proteases) are the second largest family of serine proteases in terms of number of sequences and characterized peptidases. According to MEROPS, subtilases are grouped under the S8 family of SB clan and are characterized by a catalytic triad of three amino acids, namely aspartate, histidine, and serine. They are further classified into two sub families namely subtilisins (S8A) and kexins (S8B) (Rawlings et al. 2008).

Marine microorganisms are emerging as a potential source of alkaline serine proteases and there are few reports on the extraction of alkaline serine protease from marine microbes with novel attributes like thermostability, salt tolerance and solvent stability (Makino et al. 1981; Manachini and Fortina 1998; Estrade-Badillo et al. 2003; Venugopal and Saramma 2006; Damare et al. 2006;

Bhaskar et al. 2007). Earlier, we had reported the production of an extracellular serine protease from an alkaliphilic and salt tolerant marine fungus, *Engyodontium album* which was characterized for its physiochemical and biochemical properties (Chellappan et al. 2006). Molecular weight of the purified enzyme protein was recorded approximately as 28 kDa and the enzyme showed maximum activity at pH 11 and 60°C. Activity at high temperature and highly alkaline pH suggests the suitability of the enzyme for its application in detergent industry. Here we report the cloning and sequencing of the protease gene from marine fungus *Engyodontium album* and the homology modelling of the deduced protein based on the crystal structure of proteinase K from *Tritirachium album* Limber. The three dimensional model equipped us to predict the overall structure of the catalytic domain and the interactions involved in substrate binding. Also, we discuss the molecular features of the model which may contribute to the thermal stability and broad substrate specificity of the enzyme. Based on this model, protein engineering strategies aimed at modulating substrate specificity, catalytic activity and stability could be developed.

## Materials and methods

### Strains, plasmids and media

The marine fungus, *Engyodontium album* BTMFS10 used in this study was originally isolated from marine sediments of the west coast of India (Suresh and Chandrasekaran 1999) and was available in the microbial culture collection of Department of Biotechnology, Cochin University of Science and Technology, India. The culture was maintained at 4°C on Bennet's agar medium (0.2% casein enzymatic hydrolysate, 0.1% beef extract, 0.1% yeast extract and 1% dextrose) prepared in 50% seawater. The culture was transferred to GPYS liquid media (0.1% dextrose, 0.05% peptone, 1% yeast extract, prepared in 50% seawater), incubated at 25°C for 4–7 days under static condition and the mycelial mat obtained was used for DNA isolation. For extracting the total RNA, the fungus was cultured in Domingues liquid media ($KH_2PO_4$ 15 g/l, $(NH_4)_2SO_4$ 5 g/l, $MgSO_4.7H_2O$ 1.2 g/l, $FeSO_4.7H_2O$ 0.0027 g/l, $MnSO_4.7H_2O$ 0.0016 g/l, $ZnSO_4.7H_2O$ 0.0014 g/l, $CoCl_2.6H_2O$ 0.0036 g/l, $CaCl_2.2H_2O$ 0.8 g/l, NaCl 10.0 g/l) supplemented with 1% casein following the same condition. *Escherichia coli* DH5α was used as the host organism for the construction of partial genomic DNA library and for amplifying the plasmids carrying the PCR product. The DH5α transformants were grown in Luria–Bertani medium (1% tryptone, 1% NaCl, 0.5% yeast extract) with ampicillin (50 μg/ml). The plasmid vector pMOS Blue (Amersham Pharmacia Biotech, Germany) was used for the cloning of PCR product and the plasmid vector pUC18 (Bangalore Genei, India) was used for the construction of partial genomic DNA library.

### Isolation and manipulation of nucleic acids

The mycelial mats were harvested, rinsed in sterile water, pat dried and ground to a fine powder in liquid nitrogen. High molecular weight genomic DNA was extracted from the fungus according to the CTAB method of Rogers & Bendich (Rogers and Bendich 1994) with minor modifications. Total RNA extraction was done according to the manual of TRI REAGENT™ (Sigma–Aldrich, St. Louis, USA) and RNA was stored at −70°C. All DNA manipulations were performed according to standard protocols (Sambrook et al. 1989). DNA ligations and restriction digestion reactions were all performed according to the manufacturer's recommendations. Plasmid extractions and transformation of *E. coli* with plasmid DNA were done according to Sambrook et al. (1989).

### PCR based cloning of *Eap* gene fragment

A set of degenerate primers (ASP 1F-5′ AAGTACATYGTCAAGYTCAAGGA 3′ and ASP1R-5′ GCCATSGARGTRCCMGAGATGG 3′) were designed on the basis of the conserved motifs of different fungal alkaline protease entries available in NCBI. The primers were used to amplify the *Eap* gene fragment using genomic DNA as the template. The 878 bp PCR product generated was eluted from the agarose gel using Wizard PCR Prep DNA purification kit (Promega, USA), ligated into the plasmid vector pMOS Blue and was transformed into *E. coli* DH5α. The transformants were selected on LB agar plates containing ampicillin. The recombinant plasmid designated as p*Eap1* was isolated, purified and sequenced. Identity of the *Eap* gene fragment was confirmed by comparing it with other fungal alkaline protease genes entries available in NCBI GenBank.

### Partial genomic DNA library construction and screening

To clone and characterize the full gene sequence of protease from *E. album* BTMFS10, a partial genomic library was constructed. The high molecular weight chromosomal DNA of *E. album* was digested with *Pst* 1, the fragments ranging from 1.5 kb to 5 kb size purified from the gel and ligated to the plasmid vector pUC18 pre-digested with *Pst* 1. The ligated DNA was used to transform *E. coli* DH5α and the transformants were selected by blue-white screening. Screening for clones with full gene insert was

done by colony hybridization with *Eap* gene fragment as the probe. The probe labeled with digoxigenin was detected with CSPD by chemiluminescence. The probe labeling and hybridization of the membranes were done with DIG High Prime DNA labeling and detection Starter Kit II (Roche Applied Sciences, Germany) according to the manufacturer's protocol. Two positive clones (p*Eap23* and p*Eap288*) were selected and the sizes of the insert present in both the plasmids were compared to the standard DNA markers after digestion with *Pst*1. The inserts which were of the size 1.4 and 4.5 kb, respectively, were purified and sequenced.

### Amplification of *Eap* cDNA

The gene specific primers; (*Eap*UTR F-5′ TCATCAACAGCCATCGCAGCAATAC 3′) and (*Eap*UTR R-5′ GACTAAATATGGTCGTAAGACCGATATGAATG 3′) were designed based on the 5′ and 3′ non-coding sequences of *Eap* gene obtained from the postive genomic clone p*Eap23* using GENETOOL LITE (version 1.0) software (DoubleTwist; BioTools, Inc., Alberta, Canada). *Eap* protease cDNA was amplified from *E. album* total RNA using Access RT-PCR System (Promega, USA) which yielded approximately 1350 bp product. The amplified product was gel eluted, cloned in pMOS blue and sequenced.

### Characterisation of the nucleic acid sequence and deduced amino acid sequence using various bioinformatic tools

DNA sequences obtained from the PCR reaction and the genomic library clones were assembled using the Genetool software. The homology search of DNA sequences and deduced amino acid sequences were done using the online BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/) (Altschul et al. 1997). The multiple alignment of the sequences were done with BioEdit software (Hall 1999) and the dendrogram showing the phylogenic relation of EAP with other proteases was constructed using the *MEGA* 3 program (Kumar et al. 2004). The ORF translation and the calculation of the theoretical MW of the deduced protein were done at ExPASy Proteomics server ( http://us.expasy.org/tools/pi_tool.html). Signal peptide predictions were done using SignalP 3.0 at the CBS prediction server (www.cbs.dtu.dk/services/) and the prediction of protein localization was performed with the TargetP 1.1 program at the CBS prediction server. For finding out protein motifs and domains, the following tools were used: Motif search for aminoacid patterns in the deduced protein using the ScanProsite tool in the ExPASy proteomics server (Baroch et al. 1997) and Analysis of protein domains and functional sites in EAP using InterPro Scan (Integrated Resource of Protein Families, Domains and Sites; http://www.ebi.ac.uk/interpro) (Mulder et al. 2003). Homology modelling was done using the program Swiss-PDB Viewer, SPDBV (http://www.expasy.ch/swissmod/SWISS-MODEL.html). The quality of the homology model was analysed using PROCHECK (Laskowski et al. 1993). Molecular graphics of the three dimensional models was done with VMD (http://www.ks.uiuc.edu/Research/vmd/). Prediction of calcium binding sites in the three dimensional model of EAP was performed with the GGv1.0 package (http://www.chemistry.gsu.edu/faculty/Yang/GG.htm) and the disulphide linkage prediction was done with the DISULFIND program (http://www.disulpfind.dsi.unifi.it). The number of hydrogen bonds and salt bridges were calculated using the program Accelrys Discovery Studio Visualizer 2.5 (http://accelrys.com/products/discovery-studio/visualization/discovery-studio-visualizer-25.html). The physiochemical properties like molec ular weight, amino acid composition, isoelectric point (pI), charge distribution, and the extinction coefficient (EXTCOEF) at 280 nm of the mature EAP and PRK1 were calculated using the ProtParam tool (http://www.expasy.ch/tools/protparam.html) in the ExPASy proteomics server.

## Results

An internal fragment of the gene encoding alkaline serine protease from *E. album* (*Eap*) was amplified by PCR using a set of degenerate primers. This 878 bp PCR product was cloned in pMOS vector and the recombinant plasmid (designated as p*Eap1*) was then sequenced. Homology search confirmed that the sequence belongs to the coding region of the *Eap* and has a high sequence similarity with several other fungal serine protease genes. This gene fragment was further used as the homologous probe for screening the sub genomic library which resulted in the selection of two positive clones, designated as p*Eap23* and p*Eap288*, which were isolated, purified, and sequenced. The sequences obtained when queried with the nucleotide database using the online BLASTN program (Altschul et al. 1997), showed high similarity with the known sequences of subtilases from different fungi. This information was then used to deduce the full length gene including its 5′ and 3′ noncoding regions.

The complete coding region of the *E. album* alkaline serine protease gene (*Eap*) could be obtained in the clone p*Eap23*. Using a set of gene specific primers (*Eap* UTR Forward and Reverse) designed from the 5′ and 3′ noncoding sequences of the *Eap* gene, a 1350 bp cDNA was amplified and sequenced. Comparison of *Eap* gene sequence to that of other fungal serine proteases revealed a significant similarity (96–87.5% sequence identity to *T. album* proteinase R and K, 67.4–67.5% with *Fusarium*

sp. Alp and *Lecanicillum psaliotale* Ver112, 64–63.9% with *Beauveria bassiana* bsn1 and *Cordyceps brongniartii* pr1, 62.5–63.9% with *Verticillum chlamydosporium* p1 and *Tolypocladium inflatum* prots and 65.5% with prlA of *Metarhizium anisopliae*).

The 1424 bp EAP locus contained an open reading frame of 1161 bp flanked by 93 bp UTR at the 5′ end and 102 bp in the 3′ end. The ORF is interrupted by a putative 65 bp intron, which has the characteristic features of intron sequences observed in other filamentous fungi (Gurr et al. 1988). Further, comparison of the nucleotide sequence of the genomic clone (p*Eap23*) with that of the cDNA confirmed the presence of the single intron (65 bp) in the coding region of EAP. The *Eap* gene sequence has been deposited in NCBI GenBank (http://www.ncbi.nlm.nih.gov/Genbank/) under the accession number, DQ 268654.

## The deduced aminoacid sequence of *Eap*

The *Eap* gene encoded a polypeptide consisting of 387 amino acids with a calculated molecular mass of 40.923 kDa. The putative mature protein (residues 109–387) is 279aa in length with a calculated molecular mass of 29 kDa. The deduced protein exhibited 96.0, 84.0, 62.3, 62.0, 59.8, 58.6 and 49.3% sequence identity, respectively, to proteinase R of *Tritirarchium album* (Samal et al. 1990); proteinase K of *Tritirarchium album* (Gunkel and Gassen 1989); alkaline protease of *Fusarium* (Morita et al. 1994); Pr1 of *Beauveria brongniartii* (Sheng et al. 2006); bsn1 of *Beauveria bassiana* (Joshi et al. 1995); pr1 of *Metarhizium anisopliae* (St. Leger et al. 1992) and pr1B of *Metarhizium anisopliae* (Joshi et al. 1997). A dendrogram was created from the sequence alignment of various members of serine peptidase family S8 (Fig. 1) using the program, *MEGA* version 3.1 (Kumar et al. 2004). The percentage similarity values and the dendrogram clearly showed that EAP and proteinase R of *T. album* are orthologues and also that EAP is closely similar to the *T. album* proteinase K.

The analysis of the protein sequence with TargetP v1.1 and signal SignalP 3.0 confirmed the presence of a signal peptide (21aa) at the N-terminal end of the deduced protein, indicating that EAP could be secreted. When the sequences of selected subtilases were aligned with EAP, several regions were found to be highly conserved, noted to be important in contributing to the catalytical activity and structural features of subtilases (Fig. 2). They include: (1) the consensus sequences flanking the catalytic triad (Asp39, His69, and Ser224) of the subtilisins. (2) two stretches; (Ser132–Ile133–Gly134) and (Ala158–Ala159–Gly160) which form part of the specific crevice in subtilisin within which the P1 site of the substrate side chain fits. The corresponding upper part of the crevice in proteinase K

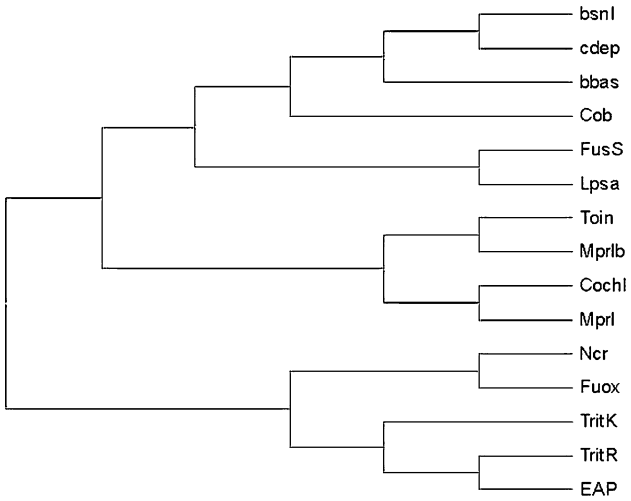


**Fig. 1** Unrooted neighbor-joining tree showing the relationship between EAP and other fungal subtilases. TritR (*Tritirarchium album* proteinase R; P23653), TritK (*Tritirarchium album* proteinase K; P06873), FusS (*Fusarium* sp. (S-19-5) ALP; AAC60571.2), Toin (*Tolypocladium inflatum* protease; AAL75579.1), Lpsa (*Lecanicillium psalliotae* ver112; Q68GV9), bsnl (*Beauveria bassiana* bsn1; AAD29255.1), Cob (*Cordyceps brongniartii* Pr1; AAR97273.1), Mprl (*Metarhizium anisopliae* pr1A; P29138), Mprlb (*Metarhizium anisopliae* pr1B; AAC49831.1), Ncr (*Neurospora crassa* probable endopeptidase K; CAD71122.1), Fuox (*Fusarium oxysporum* serine protease precursor; BAD72940.1), bbas (*Beauveria bassiana* subtilisin-like protease precursor; AAC48979.1), Cochl (*Cordyceps chlamydosporia* alkaline serine protease; CAD20578.1), cdep (*Beauveria bassiana* CDEP-1; AAK70804.1), EAP (*Engyodontium album* alkaline serine protease precursor; ABK96984.1)

(Ser129–Leu130–Gly131) forms hydrogen bonds with the P2 and P3 residues on the substrate (Kraut 1977; Betzel et al. 1986). (3) the oxyanion hole residue (Asn161 in EAP), which helps to stabilize oxyanion in a transition state in most subtilisins (Kraut 1977). (4) The cysteine residue (Cys73 in EAP) which corresponds to the free cysteine residue near the active site histidine in proteinase K has been postulated to play a role in catalysis (Betzel et al. 1986). (5) Among the internal helix like structure near active site residues, the internal helix hC and hF, which are highly conserved in subtilases, are also observed in EAP [HGTHCAGT:69–76; GTSMATPHVAGLAAY:222–236] (Siezen et al. 1991).

The motif search for amino acid patterns in EAP performed against Prosite database identified the presence of the three signature patches around the catalytic triad residues which are characteristic of subtilases. A protein domain search against the InterPro database revealed the presence of two highly conserved domains in EAP: a subtilisin N terminal domain (pfam 05922) at the sub-terminal position (between residues 40 and 108) and a Peptidase S8 domain (pfam 00082) between amino acid positions 120–377.

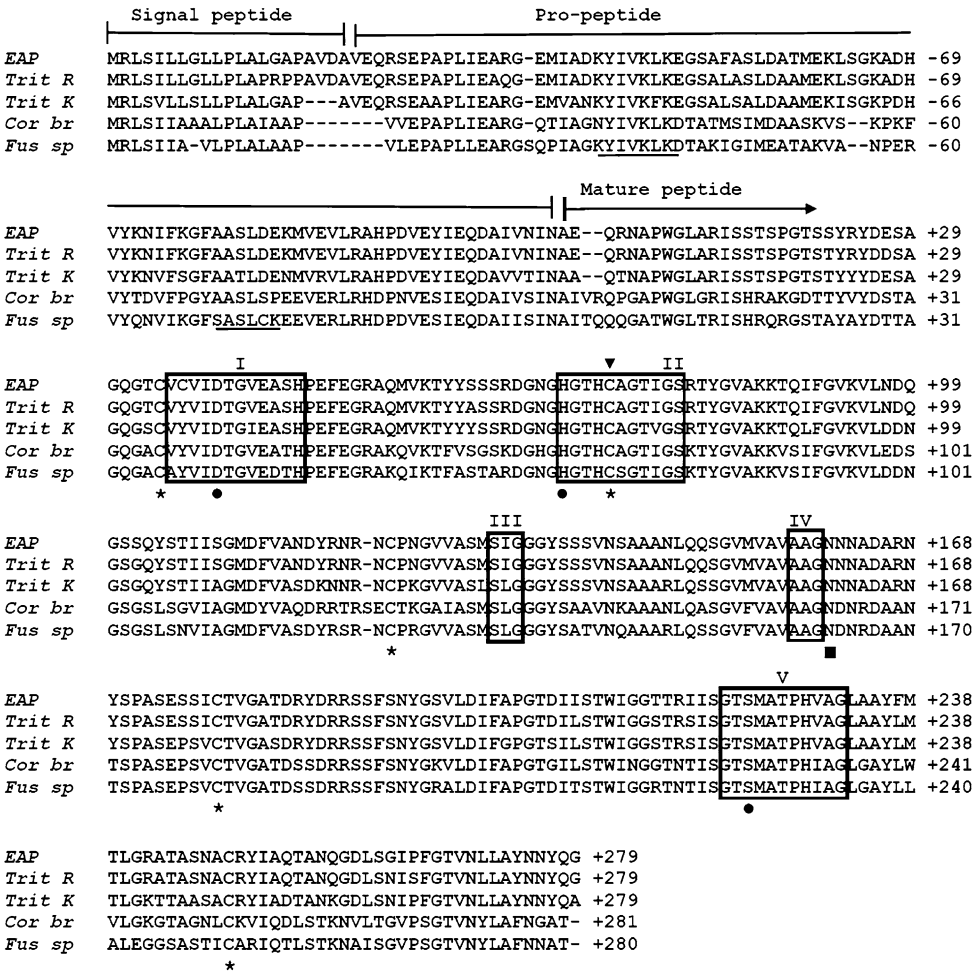


**Fig. 2** Multiple sequence alignment of the deduced aminoacid sequence of EAP (*E. album* protease; ABK96984.1) with other fungal subtilases viz., Trit. R (*T. album* proteinase R; P23653), Trit. K (*T. album* proteinase K; P06873) Cor. br (*C. Brongniartii Pr*1; AAR97273.1) Fus sp. (*Fusarium* sp. *alp*; AAC60571.2). Amino acids are numbered with respect to the first amino acid of the sequence. Gaps are introduced to optimize the alignment. The signal peptide, pro-region and starting of mature peptide are marked. The two conserved stretches of residues in the pro region are underlined. *Boxes* (I to V) indicate conserved stretches assigned to subtilisin type serine protease motifs. * Marks represent the highly conserved cysteine residues present in all the five sequences. The three *black circles* represent the catalytic triad residues of the proteases. The *small black rectangle* indicates the highly conserved asparagine residue and the *small black triangle* indicate the conserved free cysteine residue, both of which have important roles in catalysis

## Homology modelling of EAP

The sequence of EAP corresponding to the putative mature protein (residues 109–387) was used in homology modelling and was performed at the Swiss Modeller web server using DeepView. Proteinase K/PRK1; [PDB:1IC6] (Betzel et al. 1988a) was used as a structural template to model EAP, and Proteinase K complexed with the inhibitors 2-methoxysuccinyl-Ala-Ala-Pro-Ala-chloromethyl ketone; [PDB:3PRK] (Wolf et al. 1991) and N-Ac-Pro-Ala-Pro-Phe-DAla-Ala-Ala-Ala-NH; [PDB:1PFG] (Saxena et al. 1996) were the templates used to visualize the substrate binding pockets. The three dimensional model was found to be reliable with good stereochemical quality as determined by PROCHECK (Laskowski et al. 1993). Ramachandran diagram $\varphi$–$\psi$ plots for the 3D model of EAP showed that 89.5% of the residues (excluding Gly and Pro) had Phi–Psi angles in the most favorable region, 10.5% of residues in the allowed region, while none of the residues were observed in the disallowed regions of the plot or even in the generously allowed regions. Also, the main chain and side chain parameters were all within acceptable limits. The model of EAP is available in the PMDB database under the accession number PM0075214. The EAP model shows the characteristic $\alpha/\beta$ scaffold of subtilisin-like serine proteases consisting of six $\alpha$ helices, three 3/10 helices, a seven-stranded parallel $\beta$ sheet, and three-two-stranded antiparallel $\beta$ sheets. In PRK1, the active site consists of the catalytic triad Asp39, His69 and Ser224 as well as the oxyanion hole residue Asn161. When the 3D model of the active site of EAP was superimposed on the corresponding region of Proteinase K (3PRK) it was observed that the active site is absolutely conserved. Further, there were no aminoacid substitutions observed in the residues closer (within 4 A°) to the catalytic triad (Fig. 3a).

Based on sequence alignments and a detailed analysis of the three dimensional models of both PRK1 (3PRK) and EAP, following observations were made regarding the substrate binding pockets which determine the specificity and selectivity of the enzyme (Fig. 3b). In PRK1, the S2′ pocket is a hydrophobic site formed primarily by residues 192, 221 and 222. In EAP, the residues corresponding to these positions are found absolutely conserved. The S1 is a distinct large and elongated pocket in PRK1 which is

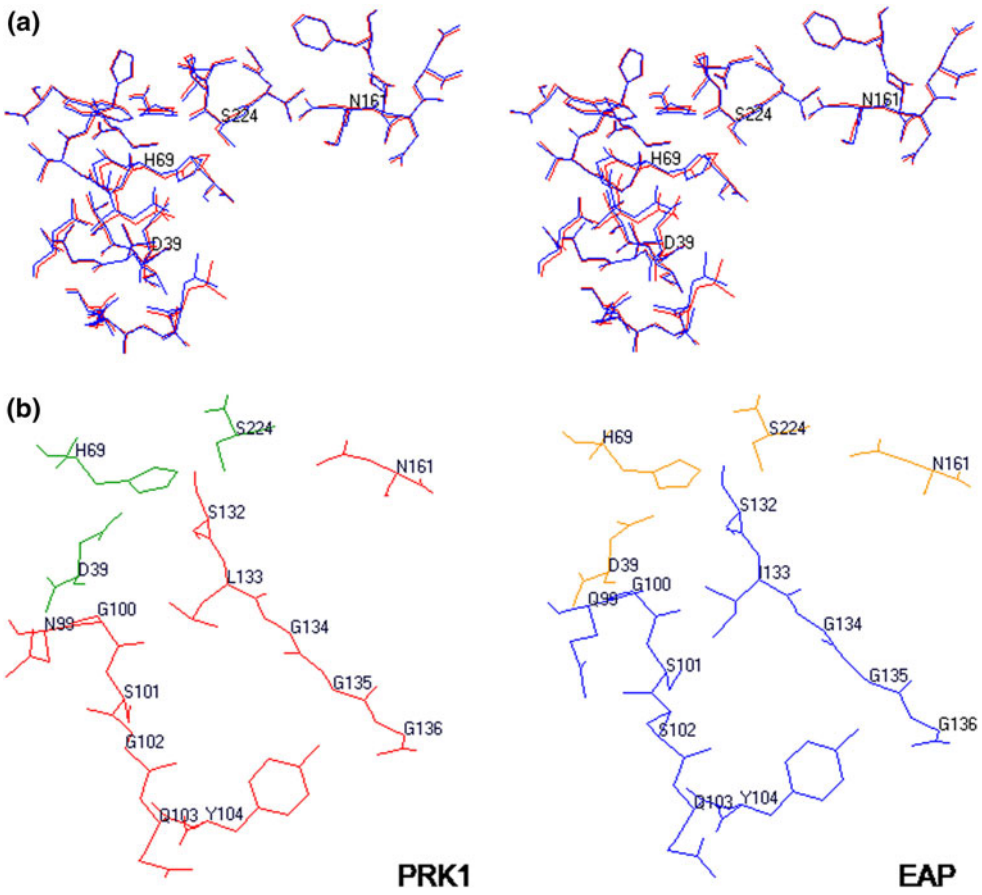


**Fig. 3** **a** Stereo image of the active site of EAP superimposed on that of PRK1. *Blue lines* represent EAP (*Darker lines* in printed version) and *Red lines* represent PRK1 (*Lighter lines* in printed version). **b** Representation of the substrate recognition site in Proteinase K and EAP

primarily formed at the side and bottom by segments 132–135 and 158–161, respectively and it is also surrounded at the rim by residue 162, at the bottom end by 169, and at the top by segment 222–225 carrying the nucleophilic residue Ser224. When the S1 pocket of EAP was compared to that of PRK1, a single amino acid substitution was noted. Leu at position 133 in PRK1 is replaced by Ile in EAP.

In PRK1, the S2 is a less distinct and smaller pocket than S1. It is bounded at one side by residue 100 and at the other side by the active site residue His69, at the bottom by both the hydrophobic residue 96 and the active site residue Asp39, at the bottom end by residue 40, and at the rim by residue 67, respectively. All these S2 site residues are absolutely conserved in EAP. The S3 site does not seem to form a pocket since the side chain of a Ser residue (at position 101 at this site) points outwards into the solvent, forming a protruding patch on the surface. The other residue which is also assumed to interact with the P3 side chain in PRK1 is at position 100. These two residues are conserved in EAP. S4 is a very distinct pocket in PRK1 which is formed between the two segments 100–104 and 132–136, and this pocket appears to have two sub sites, sub sites S4a and S4b. S4a is formed by residues 96 and 107 at the bottom, 133 at the side, and 102 at the rim, respectively while S4b has residues 104 at the side, 141 at the bottom, and 135 and 136 at the rim, respectively. The side chains of these residues determine the size of the S4 pocket. The S4 pocket was well conserved in EAP with two substitutions viz., Gly102 with Ser and Leu133 replaced by Ile.

Based on the homology model and sequence alignments of the protein, disulphide bonds and Ca binding sites were also predicted in EAP. In proteinase K, there are five cysteine residues at positions 34, 73, 123, 178 and 249 which were found perfectly conserved in EAP. In the three dimentional model of EAP, out of these five highly conserved Cys residues, four were found so close as to suspect disulphide linkage between them (Cys34 and Cys123; Cys178 and Cys249) (Fig. 4). The disulphide linkage prediction performed using the DISULFIND web server further confirmed the possibility of these two disulfide bonds in EAP (Ceroni et al. 2006).

In the crystal structure of PRK1, two $Ca^{2+}$ binding sites have been reported; the strong calcium binding site called as Ca1 and the weak calcium binding site named Ca4 (Siezen and Leunissen 1997). The Ca1 site is formed by the carbonyl groups of Asp200 and the carbonyl oxygen atoms of Pro175 and Val177 in PRK1. In EAP, the Pro175 and Val177 are replaced by Ser175 and Ile177 respectively.

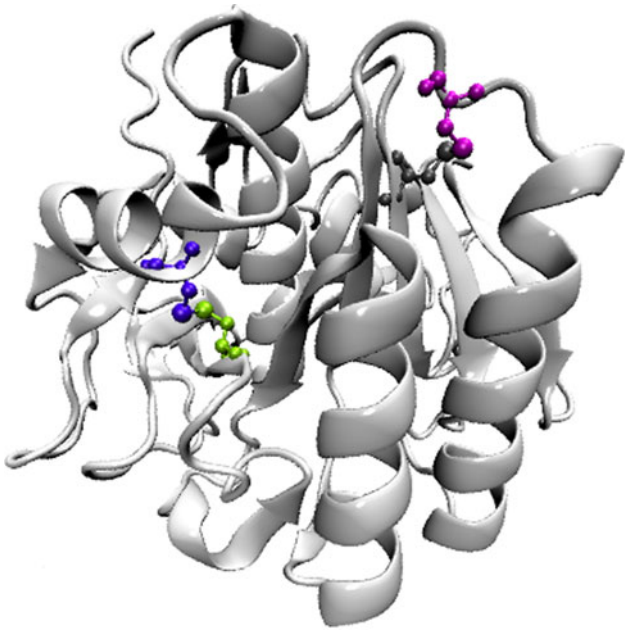

**Fig. 4** 3D model representation of EAP showing the possibility of two-disulphide bonds. The protein is represented as cartoon with *grey* colour. The cysteines predicted to form disulphide linkages are indicated as *ball* and *stick* models (Cys34-*tan* and Cys123-*majenta* (right side)); (Cys178-*green* and Cys249-*blue* (left side))

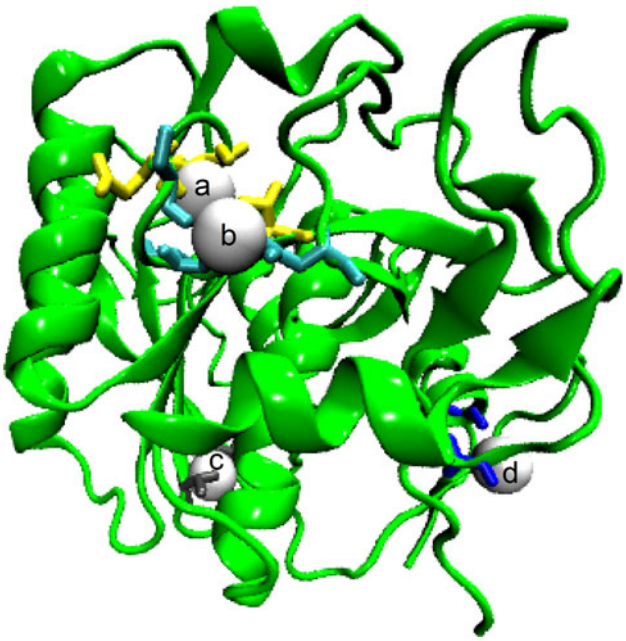


**Fig. 5** The homology model of EAP showing four putative Ca binding sites. The protein is represented as cartoon with *green* colour. The predicted Ca binding sites are shown with Ca-ions (*white spheres*) occupying the positions so as to form bonds with oxygen atoms of the residues that form the Ca binding sites. The coordinates of the four calcium ions (**a**, **b**, **c** and **d**) are represented as licorice models: **a**. (Pro171, Ala172, Glu174, Thr179-*yellow*), **b**. (Ser175, Ile177, Asp200, Asp200-*cyan*), **c**. (Gly30, Thr33, Thr33, Thr88-*tan*) and **d**. (Ala11, Ser14, Ser14, Ser22-*blue*)

The Ca4 in PRK1 is a weak binding site formed by the carbonyl atoms of Asp260 and the carbonyl oxygen atom of Thr16. In EAP, the Ca4 site ligands are absolutely conserved. So based on homology, two calcium binding sites have been suggested for EAP. However, the possibility for the presence of calcium binding sites in EAP was further analyzed by the GGv1.0 program (Deng et al. 2006). In addition to the strong Ca1 binding site (Ser175, Ile177, Asp200, Asp200), the program predicted three more calcium binding sites in the EAP structural model. viz; (Ala11, Ser14, Ser14, Ser22), (Gly30, Thr33, Thr33, Thr88), and (Pro171, Ala172, Glu174, Thr179) (Fig. 5). Earlier studies on EAP had suggested that $Ca^{2+}$ possess the role of a stabilizer promoting thermal stability during incubation of the enzyme at various temperatures, with maximal residual activity at 65°C (Chellappan 2005). This is in agreement with the prediction that EAP contain calcium binding sites in its structure. The number, nature and distribution of hydrogen bonds/salt bridges in the three dimensional model of EAP and PRK1 were analysed using Accelrys Discovery Studio Visualizer. According to the software, the number of hydrogen bonds/salt bridges in EAP is 326/14, whereas that of PRK1 is 321/15.

Considering the high sequence and structural similarities between EAP and PRK1, the physiochemical properties and the experimentally determined optimal reaction conditions of EAP were compared with that of PRK1 (Table 1). Both the enzymes share similar physiochemical properties such as molecular weight (~29 kDa), total number of aminoacid residues (279), high pI values (pI > 7.7), comparable EXTCOEF value and suceptivity to inhibition by the phenyl methane sulfonyl fluoride (PMSF). The aminoacid sequence identity of mature EAP with respect to mature PRK1 enzyme is 86%. EAP and PRK are highly active at temperatures around 60°C. For EAP, though the temperature of maximal activity was recorded at 60°C, More than 90% of the maximal activity was conserved between temperatures 45 and 65°C (Chellappan et al. 2006). The optimal pH range of EAP (9–12) is slightly narrower compared to that of PRK1 (7.5–12).

## Discussion

In this study an alkaline protease gene from a marine fungus *E. album* was cloned, sequenced and characterized. Homology comparison of the deduced amino acid sequence of *Eap* with other known proteins indicated that the gene encode an extracellular protease that belongs to the serine peptidase family S8 (Clan SB; Sub family S8A); EAP was found to be translated as a precursor protein and it is assumed that the *E. album* protease undergoes two

**Table 1** Physiochemical properties and optimal reaction conditions of EAP and PRK1

| | EAP | PRK |
|---|---|---|
| Source | Marine | Terrestrial |
| Number of amino acids (Naa) | 279 | 279 |
| Predicted isoelectric point (pI) | 7.78 | 8.25 |
| Calculated Molecular weight (mw) in kDa | 29.160 | 28.907 |
| Inhibitor | PMSF* | PMSF |
| Number of positively charged residues/number of negatively charged residues (RK/DE) | 19/18 | 20/18 |
| Predicted extinction coefficient at 280 nm | 33,725 | 36,580 |
| Sequence identity with respect to PRK (%) | 86 | 100 |
| Optimal pH range | 9–12** | 7.5–12* |
| Optimal temperature range (°C) | 45–5** | 50–60* |

*PMSF* phenyl methane sulphonyl fluoride

*Values taken from Liu et al. (2007); ** Values taken from Chellappan et al. (2006)

processing events: One after amino acid 21 to remove the signal peptide and the second after amino acid 108 to remove the propeptide and thereby releasing the mature enzyme (molecular weight ~29 kDa). The putative propeptide region next to the signal peptide functions by binding to the enzyme active site as an inhibitor and is required for proper folding and secretion of the enzyme (Tatsumi et al. 1989; Yabuta et al. 2001).

In subtilases, the substrate binding region is being described as a surface channel that can accommodate at least six amino acid residues (P4–P2′) of a substrate or an inhibitor. Both backbone and side-chain interactions are found to contribute to the substrate binding. The substrate residues P4–P1 are bound by hydrogen bonds to the two beta-strands at the substrate binding region of the enzyme to form the central strand of a three stranded anti parallel beta sheet. Interactions of the side chains of the P4–P1 and S4–S1 of the substrate and enzyme, respectively, are largely found to determine the substrate specificity. But in subtilisins, the S4 and S1 sites appears to dominate the substrate preference (Siezen and Leunissen 1997).

The single aminoacid substitution in the S1 pocket of EAP (Leu133 to Ile133) is likely to increase the hydrophobicity of the S1 pocket in EAP compared to PRK1. But the Ile residue with the beta branching does not seem to affect the S1 pocket volume as expected because its side chain seem to project outside the S1 pocket. This indicates that like Proteinase K, EAP also has a broad P1 specificity with preference for more hydrophobic residues at the P1 position of the substrate. The S2 and S3 pocket residues are absolutely conserved in EAP and PRK1 indicating that they share a common P2 and P3 specificity. In EAP, there are two substitutions in the S4 pocket residues (Gly102 with Ser and Leu133 with Ile) with regard to PRK1. In the three dimensional structure of PRK1 (and also in EAP) the residue 133 is positioned in such a way that it involves in the formation of both S1 and S4 pockets. In S4 pocket this substitution is likely to increase the hydrophobicity and simultaneously to decrease the pocket volume since the side chain of Ile133 protrudes inside the pocket. In the case of the second substitution; Gly102 with Ser, the side chain of serine protrude inside the S4 pocket which seems to decrease the size of S4a site and this change also contribute to a slight change in the hydrophobicity of the S4 pocket. Even though the size of the S4 pocket in EAP is reduced by these two substitutions, it can still accommodate bulky P4 residues like PRK1, since these two substitutions are located in the highly flexible, surface exposed loops 100–104 and 132–136, respectively (Martin et al. 1997). Especially, the segment 100–104 shows larger conformational freedom which confers on S4 pocket the capacity to accommodate a large P4 residue (Wolf et al. 1991). However, as a net effect of these substitutions, EAP may prefer more hydrophobic P1 and P4 substrate residues when compared to PRK1. The hydrophobic nature of the substrate binding region may contribute to the alkaline activity profile of the enzyme (Betzel et al. 1996).

Disulfide bridges contribute to the overall stability of a protein, and the introduction of new S–S bonds can enhance the thermal stability, as demonstrated in phage T4 lysozyme (Matsumura et al. 1989). EAP share five conserved cysteine residues with PRK1. Four of these highly conserved cysteine molecules were shown to form two disulphide bonds in PRK1 (Jany et al. 1986) and the thermostable nature of the enzyme is attributed to the presence of these two S–S bonds. By verifying the 3D model and the disulphide linkage prediction performed using the DISULFIND web server, two similar disulphide bridges were predicted in EAP. These disulphide bonds may be contributing to the thermostable nature of EAP. We had reported earlier that EAP is active over a broad range of incubation temperatures with maximal activity at 60°C (Chellappan et al. 2006).

The occurrence of calcium binding sites is a feature shared by the members of the subtilisin-like serine protease family. It has been observed that calcium binding enhances

the thermal stability of subtilisins which in turn increases their resistance to proteolysis (Betzel et al. 1988b, 1990). Also, they are essential for maintaining the correct folding and stability of the enzyme (Siezen and Leunissen 1997). In EAP, the two aminoacid replacements in the Ca1 binding site are likely to have a minor effect on calcium binding and it seems that this strong Ca1 site is retained in EAP since the carbonyl oxygen atoms of Ser175 and Ile177 can coordinate the Ca ion along with the carbonyl groups of Asp 200 in the EAP model. The Ca1 calcium binding site is found to be involved in the stabilization of strands b8 and b9. The two residues at the Ca4 calcium binding site are the same for EAP and PRK1 indicating that the Ca4 calcium binding site is retained in EAP attributing to the stabilization of the N- and C-terminal regions of the molecule. The weak Ca4 site is further strengthened by a number of hydrogen bonds and salt bridges formed between residues Arg12, Asp187 and Asp260, which are close to this calcium binding site. Prediction of the calcium binding site for the 3D model of EAP using the GGv1.0 program, led to the identification of the Ca1 site alone (which was already predicted based on crystal structure of PRK1) as a potential calcium binding site. The weak calcium binding site, Ca4, is coordinated by only two side chain ligands, and hence usually no predictions are made for such sites (Siezen and Leunissen 1997). Interestingly, apart from the strong Ca1 binding site, the program also predicted three more calcium binding sites in the EAP structural model. When the PRK1 structural model was analysed with the same program, only the Ca1 site was predicted.

There are reports of proteases containing more than two calcium binding sites, none of which correlates with the conserved calcium binding sites already reported. In the case of sphericase; a serine protease from *Bacillus sphaericus*, up to five calcium ions have been found associated with a single enzyme molecule (Almog et al. 2003), where as in the case of a marine *Vibrio* proteinase, three Ca ions were found associated with its structure, one of which was in a calcium binding site not described in other subtilases (Arnorsdottir et al. 2005). Amongst the three additional calcium binding sites predicted in EAP, the one which is formed by residues Ala11, Ser14, Ser14, Ser22 is similar to a calcium binding site reported in the marine *Vibrio* proteinase, which is formed by residues Asp11, Asp14, Gln15, Asp21 and Asn23(PRK1 numbering) (Arnorsdottir et al. 2005). In EAP, the newly observed calcium binding site is involved in stabilizing the N-terminal region of the enzyme and is strengthened by a non-conserved salt bridge; Arg24:Glu2 and two extra hydrogen bonds formed between residues Ser14–Arg12 and Asn5–Ser21. The fourth Ca binding site (Gly30, Thr33, Thr33, and Thr88) is strengthened by two additional hydrogen bonds; Thr33:Thr88 and Lys86:Gln3 and a hydrogen bond network Gly30 [Ser28, Thr33], while the fifth Ca binding site in EAP (Pro171, Ala172, Glu174, Thr179) is strengthened by the extra hydrogen bonds; Ser170:Asn142 and Ser176:Glu174 and the hydrogen bond network; Ser173 [Pro171, Tyr169, Ser175]. Considering the stabilizing effect of metal ion binding in many proteins, it might be expected that increased affinity and the number of bound metal ions should correlate with the thermostability of proteins (Arnorsdottir et al. 2005). Hence, the enhanced stability of EAP at 65°C in the presence of calcium ions may be due to the presence of the three additional calcium binding sites (with respect to PRK1) in the molecule.

Hydrogen bonds play an important role in the overall stability of the protein due to their large number and wide distribution, whereas the contribution of salt bridges to protein stability is more localized (Arnorsdottir et al. 2005). The total number of hydrogen bonds and salt bridges in EAP (326/14) is comparable with that of PRK1 (321/15). EAP and PRK1 share 249/13 hydrogen bonds/salt bridges in common. There are reports on critical ion pairs in a polypeptide chain which are responsible for temperature adaptation of the protein. These ion pairs will be unique and they bridge residues of distant region (>10 residues) (Bae and Phillips 2004). For PRK1, three critical ion pairs have been reported in its structure; Asp260:Arg12, Glu48:Arg80 and Arg147:Asp112 (Arnorsdottir et al. 2005). All of them except the last one are present in EAP structure also. The two salt bridges formed between the critical ion pair residues Arg147:Asp112 in PRK1 is not found there in the three dimensional structure of EAP as the residue Arg147 in PRK1 is replaced with Asn in EAP which disrupts the formation of these salt bridges. Instead, a unique salt bridge formed between the critical ion pairs; Glu2:Arg24 is present in EAP structure which participates in the stabilization of the N-terminus of the protein.

Among the alkaline proteases of the subtilisin family, the highly alkaliphilic enzymes contain an excess of positively charged residues on their surface (Shirai et al. 1997). Further, a direct correlation of higher optimum pH with an increase in the number of arginine residues was observed in AprM, a Bacillus serine protease (Masui et al. 1994). In AprM many of these Arg residues reside on the surface of the molecule which is suggested to enhance its stability under highly alkaline conditions. Interestingly in EAP a similar situation was observed, where thirteen out of the fifteen Arg residues are on the surface of the molecule which is assumed to enable the enzyme to adapt to highly alkaline conditions by altering their surface charge at higher pH.

Recently, it was demonstrated that the optimum pH of a protein results from two factors such as amino acid composition and the organization of the titratable groups within

the 3D structure (Alexov 2004). The relative importance of these two factors varies among the proteins. The pH optimum is mostly determined by the buried charged groups, and proteins with alkaline optimum pH will have a base/acid ratio greater than one. Here we observe that both EAP and PRK1 have base/acid ratio greater than one, which points towards the alkaline optimum pH of the enzymes. Three-dimensional structure of the protein plays an even more significant role than the sequence composition on the optimum pH and the pH of maximal stability can be calculated using the 3D structure of proteins (Alexov 2004). Present observations on hydrophobic nature of the S1 and S4 pockets, excess of Arg on the surface, and >1 value for the base/acid ratio of buried groups substantiate our previous experimental report on the alkaliphilic nature of EAP. A detailed analysis of the 3D structure of EAP will give much deeper information regarding the structural basis of the alkaliphilic nature of the enzyme.

EAP shares a very significant homology with the Proteinase K from *T. Album*. Proteinase K (PRK1) is reported to have great potential for basic and applied research due to its broad peptide cleavage activity, unusual stability over a wide range of temperatures and pH values and even in the presence of low concentrations of SDS and urea (Ebeling et al. 1974). Based on the present observation on the physiochemical, structural and sequence similarities of EAP with proteinase K, along with our previous report on the high pH and temperature optima of EAP (Chellappan et al. 2006), it is proposed that marine fungal isolate of *E. album* has the potential to serve as an alternative cum additional source of alkaline protease enzyme for various industrial and research applications. Moreover, the gene sequence and three dimensional structure of EAP reported in the present study presents ample scope for further improvement of the enzyme properties using protein engineering strategies and subsequent application of the enzyme.

**Acknowledgments** This work was supported by a research grant from the Department of Biotechnology, Government of India (grant No.: BT/PR2203/AAQ/03/109/2000). We are grateful to Dr. Kshitish Majumdhar and Dr. Rajesh Ramachandran, CCMB, India for their insightful comments and assistance in gene sequencing. We sincerely thank Dr. Shu-Qun Liu, Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, China for his valuable suggestions on three dimensional model of EAP.

## References

Alexov E (2004) Numerical calculations of the pH of maximal protein stability-the effect of the sequence composition and three dimensional structure. Eur J Biochem 271:173–185

Almog O, Gonzalez A, Klein D, Greenblatt HM, Braun S, Shoham G (2003) The 0.93 Å crystal structure of sphericase: a calcium-loaded serine protease from *Bacillus sphaericus*. J Mol Biol 332:1071–1082

Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acid Res 25:3389–3402

Arnorsdottir J, Kristjansson MM, Ficner R (2005) Crystal structure of a subtilisin-like serine proteinase from a psychrotrophic Vibrio species reveals structural aspects of cold adaptation. FEBS J 272:832–845

Bae E, Phillips GN Jr (2004) Structures and analysis of highly homologous psychrophilic, mesophilic, and thermophilic adenylate kinases. J Biol Chem 279:28202–28208

Baroch A, Bucher P, Hofmann K (1997) The PROSITE database, its status in 1997. Nucleic Acid Res 25:217–221

Betzel C, Pal G, Struck M, Jany K, Saenger W (1986) Active site geometry of proteinase K. FEBS Lett 197:105–110

Betzel C, Belleman M, Pal G, Bajorath J, Saenger W, Wilson K (1988a) X-ray and model-building studies on the specificity of the active site of proteinase K. Prot struct Funct Gen 4:157–164

Betzel Ch, Pal GP, Saenger W (1988b) Three-dimensional structure of proteinase K at 0.15 nm resolution. Eur J Biochem 178: 155–171

Betzel Ch, Teplyakov AV, Harutyunyan EH, Saenger W, Wilson KS (1990) Thermitase and proteinase K: a comparison of the refined three-dimensional structures of the native enzymes. Protein Eng 3:161–172

Betzel C, Klupsch S, Branner S, Wilson K (1996) Searching for molecular basis of biological stability of proteases at high pH. Adv Exp Med Biol 379:49–61

Bhaskar N, Sudeepa E, Rashmi H, Tamil Selvi A (2007) Partial purification and characterization of protease of *Bacillus proteolyticus* CFR3001 isolated from fish processing waste and its antibacterial activities. Bioresour Technol 98:2758–2764

Ceroni A, Passerini A, Vullo A, Frasconi P (2006) DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic Acid Res 34(Web server issue):W177–W181

Chellappan S (2005) Alkaline protease production by marine fungus *Engyodontium album* BTMFS 10. PhD Thesis, Department of Biotechnology, Cochin University of Science and Technology, India

Chellappan S, Jasmin C, Basheer SM, Elyas KK, Bhat SG, Chandrasekaran M (2006) Production, purification and partial characterization of a novel protease from marine *Engyodontium album* BTMFS10 under solid state fermentation. Process Biochem 41:956–961

Damare S, Raghukumar C, Muraleedharan UD, Raghukumar S (2006) Deep-sea fungi as a source of alkaline and cold-tolerant proteases. Enzyme Microb Technol 39:172–181

Deng H, Chen G, Yang W, Yang JennyJ (2006) Predicting calcium binding sites in proteins-A graph theory and geometry approach. Proteins 64:34–42

Ebeling W, Hennrich N, Klockow M, Metz H, Orth H, Lang H (1974) Proteinase K from *Tritirachium album* limber. Eur J Biochem 47:91–97

Estrade-Badillo C, Facundo J, Rocha M (2003) Effect of agitation rate on biomass and protease production by a marine bacterium *Vibrio harveyi* cultured in a fermentor. World J Microbiol Biotechnol 19:129–133

Gunkel F, Gassen H (1989) Proteinase K from *Tritirarchium album* limber-characterization of the chromosomal gene and expression of the cDNA in *Escherichia coli*. Eur J Biochem 179:185–194

Gurr S, Unkies S, Kinghorn J (1988) The structure and organization of nuclear gene of filamentous fungi. In: Kinghorn JR (ed) Gene structure in eukaryotic microbes. IRL Press, Oxford, pp 93–139

Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nuceic Acids Symp Ser 41:95–98

Jany KD, Lederer G, Mayer B (1986) Aminoacid sequence of proteinase K from the mold *Tritirarchium album* limber. FEBS Lett 199:139–144

Joshi L, Leger RJS, Bidochka MJ (1995) Cloning of a cuticle degrading protease from the entomopathogenic fungus, *Beauveria bassiana*. FEMS Microbiol Lett 125:211–218

Joshi L, Leger RJS, Roberts DW (1997) Isolation of a cDNA encoding a novel subtilisin like protease (Pr1B) from the entomopathogenic fungus, *Metarhizium anisopliae* using differential display- RT-PCR. Gene 197:1–8

Kraut J (1977) Serine proteases: structure and mechanism of catalysis. Annu Rev Biochem 46:331–358

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26:283–291

Liu SQ, Meng ZH, Yang JK, Fu YX, Zhang KQ (2007) Characterizing structural features of cuticle-degrading proteases from fungi by molecular modelling. BMC Struct Biol 7:33

Makino K, Koshikawa T, Nishihara T, Ichikawa T, Kondo M (1981) Studies on protease from marine bacteria. 1 Isolation of marine Pseudomonas 145-2 and purification of protease. Microbios 31:103–112

Manachini P, Fortina M (1998) Production in seawater of thermostable alkaline protease by a halotolerant strain of *Bacillus licheniformis*. Biotechnol Lett 20:565–568

Martin JR, Mulder FA, Karimi-Nejad Y, Zwan Jvd, Mariani M, Schipper D, Boelens R (1997) The solution structure of serine protease PB92 from *Bacillus alcalophilus* presents a rigid fold with a flexible substrate-binding site. Structure (London) 5:521–532

Masui A, Fujiwara N, Imanaka T (1994) Stabilization and rational design of serine protease AprM under highly alkaline and high temperature conditions. Appl Environ Microbiol 60:3579–3584

Matsumura M, Signor G, Matthews B (1989) Substantial increase of protein stability by multiple disulphide bonds. Nature 342:291–293

Morita S, Kuriyama M, Maejima K, Kitano K (1994) Cloning and nucleotide sequence of the alkaline protease gene from Fusarium sp. S-19-5 and expression in *Saccharomyces cerevisiae*. Biosci Biotechnol Biochem 58:621–626

Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Barrell D, Bateman A, Binns D, Biswas M, Bradley P, Bork P, Bucher P, Copley RR, Courcelle E, Das U, Durbin R, Falquet L, Fleischmann W, Griffiths-Jones S, Haf D, Harte N, Hulo N, Kahn D, Kanapin A, Krestyaninova M, Lopez R, Letunic I, Lonsdale D, Silventoinen V, Orchard SE, Pagni M, Peyruc D, Ponting CP, Selengut JD, Servant F, Sigrist CJA, Vaughan R, Zdobnov EM (2003) The InterPro Database, 2003 brings increased coverage and new features. Nucleic Acids Res 31:315–318

Rao MB, Taksale AM, Ghatge MS, Deshpande VV (1998) Molecular and biotechnological aspects of microbial proteases. Microbiol Mol Biol Rev 62:597–635

Rawlings N, Morton F, Kok C, Cong J, Barret A (2008) MEROPS: the peptide data base. Nucleic Acids Res 36:D320–D325

Rogers S, Bendich A (1994) Extraction of total cellular DNA from plants, algae and fungi. In: Gelvin SB, Schilperoort RA (eds) Plant molecular biology manual, 2nd edn. Kluwer, Dordrecht, pp 1–8

Samal B, Karan B, Boone T, Chen K, Stabinsky Y (1990) Isolation and characterization of the gene encoding a novel, thermostable serine proteinase from the mould *Tritirarchium album* limber. Mol Microbiol 4:1789–1792

Sambrook J, Fritish EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York

Saxena AK, Singh TP, Peters K, Fittkau S, Betzel C (1996) Strategy to design peptide inhibitors: structure of a complex of proteinase K with a designed octapeptide inhibitor N-Ac-Pro-Ala-Pro-Phe-DAla-Ala-Ala-Ala-NH2 at 2.5 A resolution. Prot Sci 5:2453–2458

Sharma J, Singh A, Kumar R, Mittal A (2006) Partial purification of an alkaline protease from a new strain of *Aspergillus oryzae* AWT 20 and its enhanced stabilization in entrapped Ca-Alginate beads. Internet J Microbiol 2:2

Sheng J, An K, Deng C, Li W, Bao X, Qiu D (2006) Cloning a cuticle-degrading serine protease gene with biologic control function from *Beauveria brongniartii* and its expression in *Escherichia coli*. Curr Microbiol 53:124–128

Shirai T, Suzuki A, Yamane T, Ashida T, Kobayashi T, Hitomi J, Ito S (1997) High-resolution crystal structure of M-protease: phylogeny aided analysis of the high alkaline adaptation mechanism. Protein Eng 10:627–634

Siezen R, Leunissen J (1997) Subtilases: the superfamily of subtilisin like serine proteases. Prot Sci 6:501–523

Siezen R, de Vos W, Leunissen J, Dijkstra B (1991) Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases. Protein Eng 4:719–737

St. Leger R, Frank D, Roberts D, Staples R (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

Suresh P, Chandrasekaran M (1999) Impact of process parameters on chitinase production by an alkalophilic marine *Beauveria bassiana* in solid state fermentation. Process Biochem 34:257–267

Tatsumi H, Ogawa Y, Murakami S, Ishida Y, Murakami K, Murakami A, Masaki A, Kawabe H, Arimura H, Motai H (1989) A full length cDNA clone for the alkaline protease from *Aspergillus oryzae*: structural analysis and expression in *Sacharomyces cerevisiae*. Mol Gene Genet 219:33–38

Venugopal M, Saramma A (2006) Characterization of alkaline protease from *Vibrio fluvialis* strain VM10 isolated from a mangrove sediment sample and its application as a laundry detergent additive. Process Biochem 41:1239–1243

Wolf WM, Bajorath J, Muller A, Raghunathan S, Singh TP, Hinrichs W, Saenger W (1991) Inhibition of proteinase K by methoxysuccinyl-Ala-Ala-Pro-Ala-chloromethyl ketone. J Biol Chem 266:17695–17699

Yabuta Y, Takagi H, Inouye M, Shinde U (2001) Folding pathway mediated by an intramolecular chaperone: propeptide release modulates activation precision of pro-subtilisin. J Biol Chem 276:44427–44434