Microbiology (2009), **155**, 2049–2057 DOI 10.1099/mic.0.027573-0

Cloning of feather-degrading minor extracellular protease from *Bacillus cereus* DCUW: dissection of the structural domains

Abhrajyoti Ghosh,[1,2]† Krishanu Chakrabarti[1] and Dhrubajyoti Chattopadhyay[1,2]

Correspondence
Dhrubajyoti Chattopadhyay
djcbcg@caluniv.ac.in

[1]Department of Biochemistry, University of Calcutta, India

[2]Dr B. C. Guha Centre for Genetic Engineering and Biotechnology, University of Calcutta, India

Received 15 January 2009
Revised 22 February 2009
Accepted 4 March 2009

Bacterial extracellular proteases play an important role in cell survival and cell–cell communication. A high-molecular-mass minor extracellular protease (Vpr) from a feather-degrading bacterium, *Bacillus cereus* DCUW, has been reported by our laboratory. In the present study, we cloned and expressed Vpr in *Escherichia coli*. Complete nucleotide sequencing of this gene predicted that the protease is a member of the serine protease family, and SMART domain analysis revealed that the protease consists of an N-terminal signal sequence for secretion, a subtilisin_N sequence that is a signature for N-terminal processing, a catalytic S_8 peptidase domain, and finally a long C-terminal protease-associated (PA) region containing nine intrinsically disordered subdomains. Four truncated constructs of the Vpr protease were cloned and expressed in *E. coli*. We found that the catalytic domain (amino acid residues 172–583) is sufficient for protease activity. Maturation of the Vpr protease needed both N-terminal and C-terminal processing. We have demonstrated that the oligomerization property is associated with the C-terminal protease-associated domain and also shown that the substrate-binding specificity to raw feather resides in this domain.

INTRODUCTION

Secretory proteins in bacteria are known to execute several very important 'remote-control' functions, such as nutrient cycling and utilization, cell-to-cell communication and detoxification of the extracellular environment, and they may also behave as weapons against potential competitors. Most of these secretory proteins are synthesized as a preproprotein with an N-terminal signal peptide, which is required to target these proteins to the preprotein translocase in the membrane and initiate the translocation process (Pugsley, 1993). Membrane-bound type I signal peptidases (Palacin *et al.*, 2002) remove this signal peptide to release the mature secretory protein from the *trans* side of the membrane either during or shortly after translocation (Dalbey *et al.*, 1997). Statistical studies of sequences surrounding the signal peptidase cleavage site led to the formulation of the −1, −3 or Ala-X-Ala rule, defining the preferred residues (i.e. Ala) at the −1 and −3 positions relative to the cleavage site as critical determinants for signal peptide recognition and cleavage (von Heijne, 1983, 1985).

Many bacterial proteolytic enzymes are synthesized as inactive precursors, or zymogens, to prevent unwanted protein degradation and to enable spatial and temporal regulation of proteolytic activity (Khan & James, 1998). Biochemical studies of the activation mechanism of individual proteases have provided insights into their physiological functions. In the *Bacillus* system, every extracellular protease is found to be synthesized as a preproenzyme in the cytoplasm and is processed to a mature enzyme in the extracellular milieu. The activation mechanism of subtilisin from *Bacillus subtilis* has been well studied. The zymogen form of subtilisin is converted to an active form via intramolecular auto-processing (Ohta & Inouye, 1990) in the extracellular milieu (Inouye, 1991). The prepeptide has been shown to guide the proper folding of subtilisin *in vivo* and *in vitro* (Strausberg *et al.*, 1993). Most extracellular proteases in the *Bacillus* system seem to be activated by auto-processing. The limited proteolysis actually dictates the auto-processing among proteases in the extracellular medium. Moreover, studies of the activation of different secreted enzymes from their pre-enzyme forms have shown that proteases have a definite role in

†Present address: Max Plank Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse, D-35043 Marburg, Germany.

Abbreviation: DLS, dynamic light scattering.

The GenBank/EMBL/DDBJ accession number for the sequence of the *vpr* gene is EU626488.

the post-translocational activation of different secretory enzymes (Sarvas *et al.*, 2004).

The Gram-positive, spore-forming bacterium *Bacillus cereus* contains a total of 51 protease-encoding CDSs in its genome (Ivanova *et al.*, 2003). Under different nutrient conditions, one or more of these proteases are induced. Among the major extracellular proteases, Epr, AprX and Bpr are absent from the genome of *B. cereus* (Ivanova *et al.*, 2003). Therefore, extracellular protein degradation is mainly brought about by neutral protease (Npr), Zn-metalloproteases, alkaline serine proteases, etc. Minor extracellular protease (Vpr) is present in the genome in more than one copy but the conditions under which this enzyme gets induced were not known at the date of this study. We have identified a feather-degrading protease from *B. cereus* DCUW (Ghosh *et al.*, 2007, 2008) and characterized it as Vpr. We found that in feather-supplemented medium only Vpr was induced as the sole extracellular high-molecular-mass protease and also demonstrated that thiol reductase was assisting Vpr during the feather-degradation process (Ghosh *et al.*, 2008). Vpr was found to use fibrin, collagen, gelatin, casein and also albumin as protein substrates.

Minor extracellular protease Vpr has been reported in *B. subtilis* (Park *et al.*, 2002; Sloma *et al.*, 1991) and *Bacillus licheniformis* (Ageitos *et al.*, 2007), and it has been shown that Vpr has major roles in subtilisin (lantibiotic) (Corvey *et al.*, 2003) processing, competence and sporulation factor (CSF) processing, and probably in biofilm formation and quorum sensing (Lanigan-Gerdes *et al.*, 2007). Moreover in *B. subtilis*, Vpr was identified as a fibrinolytic enzyme (Park *et al.*, 2002) and in *B. licheniformis*, Vpr was reported as one of the milk-clotting proteases (Ageitos *et al.*, 2007). Because of its unusual properties of forming large protease complexes and progressive autoprocessivity, Vpr has received increased attention over the last few years. Still we do not have enough data from molecular investigations that could explain the various activities of Vpr. Why a bacterium synthesizes such a large extracellular proteolytic enzyme was not known for a long time. Recent evidence from B. C. Park's and B. A. Lazazzera's lab further supported the importance of the large size and the long C-terminal domain in Vpr (Kho *et al.*, 2005; Lanigan-Gerdes *et al.*, 2007). Further structural insights are still important to understand details of auto-processing and substrate binding domain(s) of minor extracellular protease Vpr.

In the present study, we were interested to characterize the structure–function relationship of Vpr protein from *B. cereus* DCUW. We have cloned the *vpr* gene and expressed the functionally active Vpr protease in *Escherichia coli*. Using an antibody raised against recombinant Vpr in rabbit, we have shown the *in vitro* autoprocessivity of overexpressed full-length Vpr protein and also of the secreted Vpr of *B. cereus* DCUW grown in feather-supplemented medium. We have constructed four deletion mutants of the Vpr protein, and tried to determine the minimal domain required to show protease activity and also the importance of the long C-terminal domain.

METHODS

Micro-organisms and growth conditions. *Bacillus cereus* DCUW used in this study, and the feather-supplemented medium (FSM) used for growth, have been described previously (Ghosh *et al.*, 2008). The gene cloning and expression of recombinant genes was done using *Escherichia coli* XL1-Blue and BL21(DE3), respectively. All recombinant strains were aerobically cultivated in LB appropriately supplemented with ampicillin (100 μg ml^{-1}) at 37 °C.

Cloning, expression and sequencing of *vpr*. The total RNA from FSM-grown *B. cereus* DCUW was used as a template for reverse transcriptase using *vpr*-specific reverse primer VprR (Ghosh *et al.*, 2008). The PCR was performed using c-DNA as template and VprF and VprR (Table 1) as primers. The PCR protocol was as follows: denaturation at 94 °C for 3 min, followed by 35 cycles of 40 s at 94 °C, 45 s at 50 °C, and 2.5 min at 72 °C using High Fidelity DNA polymerase (USB) in a thermal cycler (model 9600; Applied Biosystems). The 2754 bp PCR product obtained was double digested with *Nde*I/*Bam*HI restriction enzymes and ligated with the pET20b plasmid double-digested with the same restriction enzymes using T4 DNA ligase (Fermentas) at 16 °C overnight and then used for transformation of XL1-Blue. Positive clones were confirmed by restriction digestion and also by colony-PCR using appropriate primers. The recombinant plasmid containing the *vpr* gene insert (pET-20b-Vpr) was then isolated from the XL1-Blue strain.

DNA isolated from the pET-20b-Vpr construct was used to sequence the full *vpr* gene. Mini-prep (Qiagen) columns were used to isolate small-scale pET-20b-Vpr from an overnight culture of XL1-Blue transformed cells. Sequencing was performed by the PCR-walking technique in a Genetic Analyzer 3130 (ABI) using the Big Dye Terminator cycle sequencing kit V 3.1 (ABI); primers used for sequencing were T7F, VprF, DJCF, V3F, V1R, DJCR, VprR and T7R (Table 1).

The recombinant pET-20b-Vpr was transformed into BL21(DE3) for expression analysis. The transformed cells were cultured at 37 °C overnight in 2× LB medium supplemented with ampicillin (100 μg ml^{-1}) and then transferred to fresh medium with a ratio of 1 : 100. When the OD_{600} reached 0.4, IPTG was added to a final concentration of 1 mM. After induction for 4 h, cells were harvested by centrifugation. Cell pellets were then suspended in a lysis buffer

Table 1. Sequence of the primers used in the present study

Primer	Sequence
VprF	5′-ggaattccatatgaaaaaaactacatctacac-3′
VprR	5′-cgggatccttatttcacttctaaattaacag-3′
DJCF	5′-cgcttgatggaaaaggtatgaaag-3′
DJCR	5′-gaattaataatttcaagtcc-3′
V1F	5′-ggaattccatatggaatcactacaaaagg-3′
V2F	5′-ggaattccatatgatatcagtaattgtagaattac-3′
V3F	5′-ggaattccatatgaacgaagaagcacc-3′
V4F	5′-ggaattccatatgttaattaacattcc-3′
V1R	5′-cgggatccgttaattaaaccagatcc-3′
T7F	5′-taatacgactcactataggg-3′
T7R	5′-gctagttattgctcagcgg-3′

(50 mM Tris/HCl, 250 mM NaCl, pH 8.0) and Sarkosyl was added to a final concentration of 0.7 % (to increase the solubility of overexpressed Vpr, as without Sarkosyl Vpr mostly remained in the inclusion bodies). The suspended cells were then disrupted by sonication. The supernatant after centrifugation at 12 000 r.p.m. (17 226 ***g***) was concentrated using a centrifugal concentrator membrane filter (Amicon).

Cloning and expression of four deletion mutants of Vpr protein. DNA sequencing analysis revealed that a single ORF was present in a 2.754 kb DNA fragment, which corresponds to the minor extracellular protease. The ORF of *vpr* coded for a protein of 917 aa with a putative mass of 98.843 kDa. Based on the bioinformatics analysis using the Simple Modular Architecture Research Tool (SMART) (Letunic *et al.*, 2006), InterProScan (Quevillon *et al.*, 2005) and Hits (Pagni *et al.*, 2004), four deletion mutants of the *vpr* gene were constructed. The four deletion mtants are (i) Vp1 (25–586 aa) with no N-terminal signal sequence or C-terminal sequence; (ii) Vp2 (61–586 aa) with subtilisin_N and peptidase_S8 domains; (iii) Vp3 (165–586 aa) with only the peptidase_S8 domain; and (iv) Vp4 (583–917 aa) with only the C-terminal protease-associated (PA) domain. All the deletion mutants were constructed by the PCR amplification technique with pET-20b-Vpr as the template. The primers used for amplification of Vp1, Vp2, Vp3 and Vp4 were V1F, V2F, V3F, V4F, V1R and VprR respectively (Table 1).

The expression of Vp1, Vp2, Vp3 and Vp4 deletion mutants was carried out using a similar protocol as used for full-length *vpr* from pET-20b-Vpr. All the deletion mutants were checked by SDS-PAGE and gelatin zymography (Lantz & Ciborowski, 1994).

Purification of overexpressed Vpr and the deletion mutants. The concentrated soluble Vpr was purified by affinity chromatography using an α-casein agarose column (5 × 2.5 cm) (Sigma). The protein was loaded on the column, which was equilibrated with 50 mM Tris/HCl, pH 8.0, contaning 2 mM $CaCl_2$ (buffer A). The bound protein was eluted from the column by two steps using buffer A plus 1 M NaCl (buffer B), followed by 25 % (v/v) 2-propanol in buffer B, each at a flow rate of 0.5 ml min^{-1}. Active fractions were pooled and concentrated by using Centricon filters with a 5 kDa cut-off (Amicon). The concentrated fraction was loaded onto a Superdex 200/300 gel filtration column equilibrated with 50 mM Tris/HCl, pH 8.0, and the protein was eluted with the same buffer containing 0.1 M NaCl. Fractions of 1 ml were collected at a flow rate of 0.5 ml min^{-1}. The active fractions were pooled, concentrated and checked by SDS-PAGE.

Vp1, Vp2 and Vp3 were solubilized and purified using a similar protocol as followed in the case of full-length Vpr protein. Vp4, the C-terminal fragment, was found to be highly insoluble, even in the presence of detergents. Therefore, in the case of *E. coli* overexpressing Vp4, inclusion bodies were purified and then resuspended in 50 mM Tris/HCl, pH 8.0, containing 8 M urea. The resuspended protein was then dialysed overnight at 4 °C against 50 mM Tris/HCl, pH 8.0, containing 2 M urea. The dialysed protein was concentrated with a Centricon filter (10 kDa cut-off) and then purified as described above.

***In vitro* renaturation of Vpr and Vp1–4.** The purification strategy involving Sarkosyl-mediated protein solubilization of the recombinant proteins resulted in significant loss of protein due to auto-processing. A change in strategy was needed to keep the recombinant proteins intact and we decided to work with the inclusion bodies followed by denaturation and renaturation. Vpr was purified from the inclusion bodies using 50 mM Tris/HCl (pH 8.0) containing 8 M urea. The dissolved protein was then loaded onto DEAE-Sephadex, pre-equilibrated with the same buffer. The elution was carried out using the same buffer containing 0.1–1 M NaCl. After SDS-PAGE analysis, fractions containing the enzymic activity were pooled, concentrated and finally loaded onto an FPLC Superdex 200/300 column, equilibrated with 50 mM Tris/HCl (pH 8.0) containing 8 M urea, and eluted with the same buffer containing 0.1 M NaCl. Fractions of 1 ml were collected at a flow rate of 0.5 ml min^{-1}. Vpr-containing fractions were detected by SDS-PAGE and then pooled and concentrated for further analysis.

Renaturation of full-length Vpr purified from the gel filtration was performed by a step-wise dialysis procedure against 50 mM Tris/HCl (pH 8.0) containing 0.1 M NaCl and a variable amount of urea. The protein was dialysed against 200 volumes of buffer containing 6 M urea at 4 °C for 3 h. The initial buffer was then sequentially replaced every 2 h with buffer containing decreasing amounts of urea from 4 to 1 M. Finally, urea was completely removed from the buffer. After completion of the renaturation process, protein was TCA-precipitated and subjected to SDS-PAGE followed by immunoblotting with anti-Vpr antibody. All the deletion mutants were denatured under identical conditions and their auto-processing activities were checked by Western blotting.

SDS-PAGE and zymography. SDS-PAGE was performed on a vertical slab gel containing 10 % (w/v) polyacrylamide, by the method of Laemmli (1970).

Zymography was performed using 10 % polyacrylamide slab gels containing SDS and 0.2 % of the substrate (gelatin, collagen, fibrin or casein) as a co-polymerized substrate, as described by Heussen & Dowdle (1980). The activity band was observed as a clear colourless area depleted of substrate in the gel against the blue background.

Development of antisera against Vpr in rabbit. Purified Vpr was used to raise polyclonal antibodies in rabbit. The collected antisera were used to follow the auto-processing behaviour of the protease.

Western blot analysis. To analyse auto-processing of overexpressed Vpr, proteins were first separated using SDS-PAGE and then transferred onto PVDF membrane. The membrane was then immersed in a blocking TBST buffer (20 mM Tris/HCl, pH 7.6, 0.1369 M NaCl and 0.1 % Triton X-100) containing 5 % skim milk at room temperature for 1.5 h and then incubated overnight at 4 °C with anti-Vpr antibodies (1/10 000). The membrane was then washed several times with TBS and TBST buffers, and then incubated with horseradish peroxidase-conjugated secondary antibody for 1 h. After washes in TBST, detection was performed by an enhanced chemiluminescence (ECL) detection system (Pierce).

Dynamic light scattering (DLS). The aggregation state of the protein was evaluated by DLS (Malvern Instruments). For DLS analysis, purified Vp4, eluted from a gel filtration column, was used at 1.5 mg ml^{-1} concentration. Different concentrations of SDS (0.02–0.1 %) were used to check the oligomeric status of the protein. Since DLS measures the size distribution of the protein molecules in the solution, it is sensitive to variations in particle size and interactions of protein molecules in solution (Schmitz, 1990). Thus, protein aggregation is recognized by an increase in hydrodynamic radius (R_h).

Feather keratin binding assay. The binding capacities to feather keratin of the Vp1, Vp2, Vp3 and Vp4 proteins containing the various regions of full-length Vpr were analysed by a batch method followed by SDS-PAGE. The batch method was modified from that reported by Itoi *et al.* (2006).

One millilitre of ice-chilled buffer I (50 mM Tris/HCl, pH 8.0) was added to 20 mg feather keratin powder in a microfuge tube and equilibrated at 4 °C for 60 min in a cell-mixer. Then the mixture was centrifuged at 13 000 r.p.m. (20 217 ***g***) for 10 min and the supernatant was discarded. Fifty microlitres of the purified protein solution

(2.5 μg) to be investigated was added to 1 ml of fresh buffer I and kept at 4 °C for 180 min while mixed. BSA was used as negative control. After incubation, the mixture was centrifuged as before and the supernatant was collected for further analysis. The supernatants were then precipitated using 20 % TCA, washed with 5 % TCA and 70 % ethanol, dried, and finally resuspended in 20 μl 50 mM Tris/HCl, pH 8.0. The protein samples were then subjected to 10 % SDS-PAGE analysis and the band patterns of the gel were compared by using Quantity One software (Bio-Rad).

Various concentrations of NaCl and detergents were added to the microfuge tubes containing keratin-bound protein in the presence of 1 ml buffer I to find the conditions needed to release the bound protein from the feather keratin powder. The elution was attempted by incubating with salts and detergents of different concentrations in a cell-mixer, with the protein-bound keratin, at 4 °C for 60 min. After incubation, supernatants were collected and precipitated using TCA, as described above.

RESULTS

Cloning of the *vpr* gene and expression and purification of the Vpr protein

The *vpr* gene was cloned from the RT-PCR product of *B. cereus* DCUW. We introduced the 2754 kb RT-PCR product into the pET20b *E. coli* expression vector and then the recombinant protein was induced by addition of IPTG. An approximately 100 kDa preproprotein was expressed (Fig. 1a), which was mainly localized in inclusion bodies. The proteolytic activity of overexpressed Vpr was checked on zymography gels using fibrin, collagen, gelatin and casein as substrates. The full-length Vpr was purified to homogeneity using α-casein agarose affinity chromatography and Superdex 200/300 gel filtration chromatography (FPLC) (Fig. 1a). We found Vpr activity in the void volume of gel filtration chromatography, indicating large protease complex formation by Vpr (Fig. 1b).

Full-length sequencing of *vpr* gene and domain analysis

We sequenced the full-length *vpr* gene from *B. cereus* DCUW using the PCR-walking technique. The full-length sequence was deposited in GenBank under accession number EU626488.

Detailed analysis of the domains present in full-length *vpr* revealed the presence of pre- and pro-sequences in the 100 kDa protein. SMART domain analysis revealed that the initial amino acids 1–25 (nt 1–75) encode a putative N-terminal extracellular signal sequence in Vpr. Similarly amino acids 62–158 (nt 184–474) encode a domain similar to subtilisin_N, amino acids 172–583 (nt 514–1749) encode a peptidase_S8 domain, and finally amino acids 606–917 (nt 1818–2754) encode a long Vpr-specific protease-associated (PA) domain (Fig. 2). A combinatorial homology search using the SSEARCH (http://www-btls.jst.go.jp/cgi-bin/Tools/SSEARCH/index.cgi), FASTA (http://www.ebi.ac.uk/fasta/) and BLAST programs revealed that the minor extracellular protease from *B. cereus* DCUW resembled Vpr from *B. cereus* strains ATCC 14579, ATCC 10987 and E33L, and *B. thuringensis* serovar konkukian strain 97-27. The C-terminal domain has been predicted to be involved in substrate binding in the extracellular milieu. Four types of secondary-structure predictions were carried out for the region by using the SCRATCH Protein Predictor (SSPro) (http://scratch.proteomics.ics.uci.edu/), APSSP (Advanced Protein Secondary Structure Prediction Server; http://imtech.res.in/raghava/apssp/), GOR (Garnier *et al.*, 1996; http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html), and HNN Secondary Structure Prediction method (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html). Their combined results showed that the C-terminal PA domain, which contains nine intrinsically disordered regions, was composed of

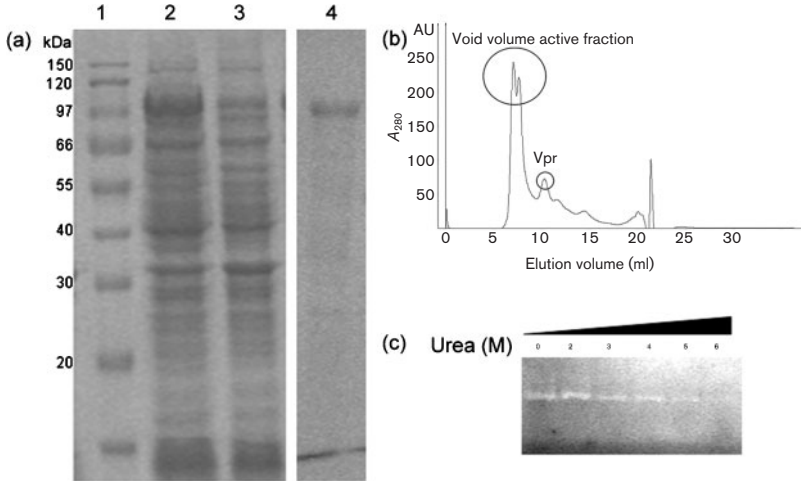

Fig. 1. Expression and characterization of minor extracellular protease Vpr. (a) Overexpression of Vpr in *E. coli* BL21(DE3) and purified Vpr by 10 % SDS-PAGE. Lane 1, molecular mass markers; lanes 2 and 3, induced and uninduced Vpr-pET20b construct in *E. coli* BL21(DE3) respectively; lane 4, purified Vpr from Superdex 200/300 FPLC column. (b) Gel filtration profile of Vpr in an FPLC Superdex 200/300 column. Active fraction in the void volume and the pure Vpr fractions are indicated by large and small circles, respectively. (c) Effect of urea on proteolytic activity of purified Vpr (recombinant) using gelatin zymography. The activity bands represent the residual activity on zymography gel after treatment with increasing concentrations of urea.

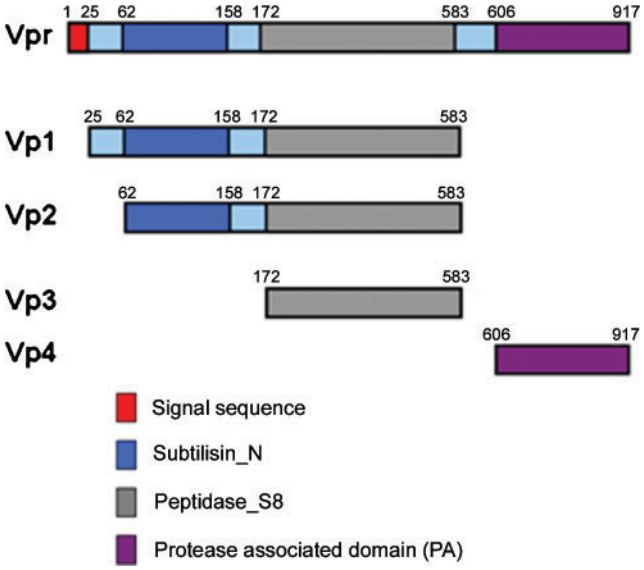

Fig. 2. Physical map representing the different domains present in Vpr using SMART software. Deduced amino acid sequence representing the entire ORF was used for the analysis. Different colours represent different domains as described in the figure.

32.24 % extended strands (Ee), which was higher compared to other domains in Vpr (data not shown) . The function of the long C-terminal domain was not known previously. The present analysis suggests that this C-terminal domain is probably responsible for the substrate binding via protein–protein interactions. Previously we have shown that this protease could degrade feather residues (Ghosh *et al.*, 2008). In the present study, the feather keratin-binding assay was designed to check whether the C-terminal domain specifically binds to feather. We also checked the aggregation pattern of the C-terminal region using a purified Vp4 mutant.

Characterization of Vpr

The overexpressed Vpr in *E. coli* was found to be biologically active. The majority of protein was obtained from inclusion bodies. When resuspended and sonicated in the presence of Sarkosyl, Vpr was found to be in soluble form, and after overnight dialysis it gave activity bands on fibrin, collagen, gelatin, and casein zymography gels (Table 2). Purified recombinant Vpr showed a major activity band at ~100 kDa, but weaker activity bands were also evident above and below the major band (data not shown). The recombinant Vpr sustained activity on zymography gels in the presence of urea up to 5 M. The protease activity was drastically reduced at urea concentrations of 6 M and higher, probably due to complete denaturation of the protein (Fig. 1c).

Expression and characterization of deletion mutants

Overexpressed deletion mutants Vp1, Vp2 and Vp3 were found to be proteolytically active. We found that Vp3, which contained only the peptidase_S8 domain, was the minimum domain in Vpr that could show protease activity. Moreover, deletion of this C-terminal protease-associated domain in Vpr resulted in reduction of the range of substrate recognition found in our study. None of Vp1, Vp2 and Vp3 was capable of degrading fibrin or collagen but they could degrade non-specific substrates like gelatin or casein (Table 2). This observation gave further insight into the domain responsible for binding specific substrates in Vpr. Furthermore, the Vp4 mutant, which contains the C-terminal protease-associated domain, was found to be proteolytically inactive, as expected from the domain analysis of Vpr.

This C-terminal domain was composed of several intrinsically disordered regions. SMART software predicted nine small intrinsically disordered regions of amino acid span 7–25 in the C-terminal domain. Vp4 was eluted from the gel filtration column as a sharp peak corresponding to a well-defined protein species with an apparent molecular mass of 37 kDa. The DLS analysis with this purified fraction showed that the hydrodynamic radius was 342 nm, which was probably due to formation of soluble aggregates that resulted from intermolecular interactions between Vp4 molecules. To check this further, the purified Vp4 was treated with different concentrations of SDS (0.02–0.1 %) and analysed by DLS. The hydrodynamic radius was found to decrease with increasing SDS concentrations; the minimum diameter (11.7 nm) was found at 0.08 % SDS,

Table 2. Results of SDS-PAGE zymography using full-length Vpr and deletion mutants Vp1–4

Casein, gelatin, fibrin and collagen were the substrates used during the zymography analysis. Proteolytic activity was measured following the protocol of Anson (1938) with some modifications. The activity unit in all cases is μmol (IU) (mg protein)$^{-1}$. SD was calculated from five independent experiments.

Substrate	Vpr	Vp1	Vp2	Vp3	Vp4
Casein	8.00 ± 0.03	7.84 ± 0.02	8.10 ± 0.02	7.65 ± 0.01	0.00 ± 0.01
Gelatin	7.50 ± 0.03	7.24 ± 0.01	7.50 ± 0.01	6.45 ± 0.03	0.00 ± 0.01
Fibrin	7.40 ± 0.01	0.00 ± 0.02	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01
Collagen	6.60 ± 0.03	0.00 ± 0.01	0.00 ± 0.03	0.00 ± 0.05	0.00 ± 0.01
Feather meal	2.55 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.01

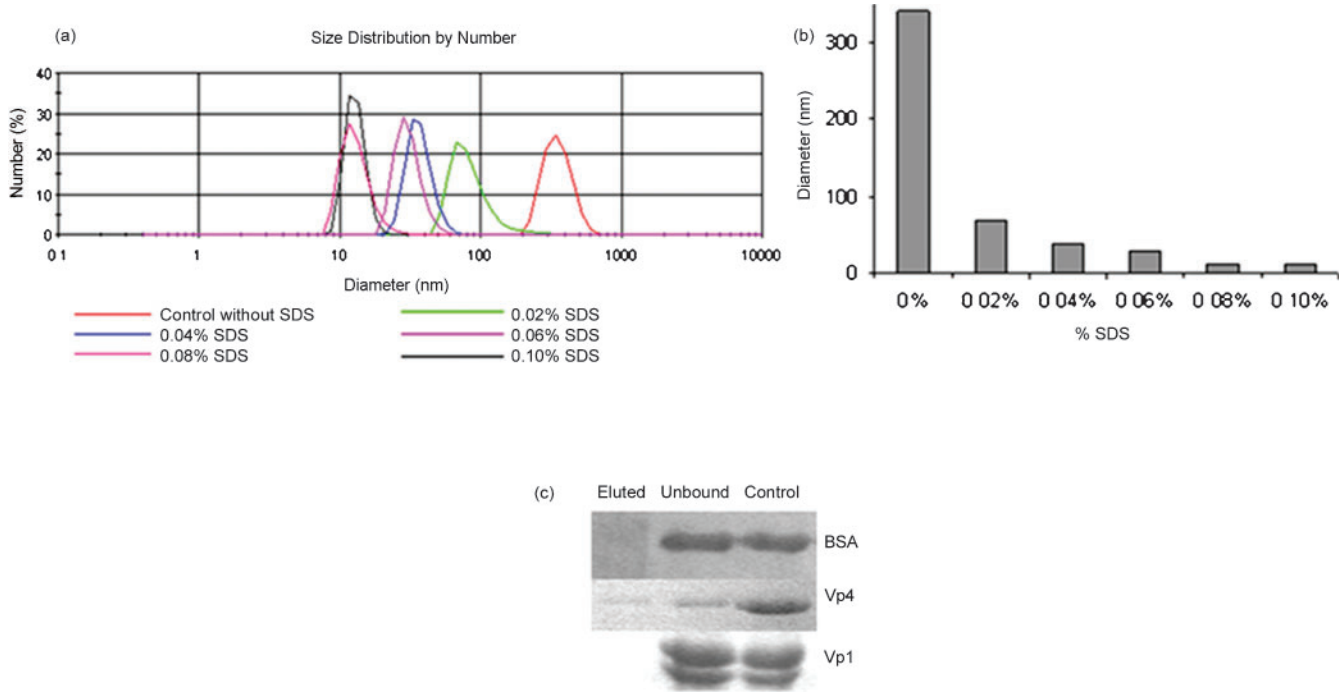

Fig. 3. Oligomerization potential and feather binding of mutant Vp4. (a) DLS experiment with purified Vp4 in the presence of 0.02–0.1 % SDS. A decrease in the hydrodynamic radii with increasing SDS concentrations indicates oligomer formation in purified Vp4. (b) Bar diagram showing the diameter of the purified Vp4 protein in the presence of SDS. The minimum diameter of 11.7 nm was found at 0.08 % SDS. No further decrease in the diameter was observed with higher SDS concentrations. (c) Feather binding of the purified Vp4, with Vp1 and BSA as control. The binding assays were carried out as described in Methods. The protein concentration of all the three proteins in the assay was 2.5 μg. Elution of the feather-bound protein fraction was carried out using 1 M NaCl. Lanes: Control (where proteins were incubated without any protein), Unbound (where protein bands represented the fraction of proteins that remained unbound to the feather after incubation), and Eluted (where protein bands represented the fraction of proteins that remained bound to feather during the assay and eluted when subjected to a treatment with 1 M NaCl).

and remained the same even in the presence of higher concentrations of SDS (Fig. 3a, b). The diameter of 11.7 nm for a 37 kDa protein is still greater than for a globular form and Vp4 is assumed to have an extended conformation in solution in the absence of its 'partner molecule'. Probably interaction with the appropriate partner molecule results in a more compact conformation as described for some other proteins having intrinsically disordered domains (Dunker *et al.*, 2001).

The feather keratin-binding assay proved to be useful to search for the substrate-binding domain in Vpr. We found that the Vp4 mutant was capable of interacting with feather keratin, and the interaction was noncovalent in nature. No detergent (SDS, Triton X-100, NP-40) was found to release bound Vp4 from the feather keratin powder in solution. Results showed that only in the presence of 1 M NaCl was Vp4 partially released from the surface of the feather keratin (Fig. 3c). Both the control BSA and Vp1 showed no feather-binding activity (Fig. 3c); therefore, we conclude that the C-terminal domain has the ability to bind specific substrates like feather keratin. This probably explains why C-terminal deletion mutants Vp1, Vp2 and Vp3 have reduced substrate recognition and binding specificity.

In vitro auto-processing of Vpr and Vp1–4

In our previous work on feather-degrading protease (Ghosh *et al.*, 2008), we found 70 ± 10 kDa protein bands during purification studies. Moreover, during the native protein purification, we found that several smaller fragments (both proteolytically active and inactive) were generated from the purified protease. In the present work, we were basically interested to see whether the recombinant Vpr in *E. coli* is capable of showing similar auto-processing. To ensure the auto-processing we performed renaturation and *in vitro* auto-processing experiments.

Since the proteins were purified under denaturing conditions, we tried to renature all the deletion mutants (Vp1–4) and full-length Vpr. After performing step-wise dialysis, all proteins were analysed by Western blotting using anti-Vpr antibody. We found that all the proteolytically active proteins (Vpr and Vp1–3) were capable of auto-processing

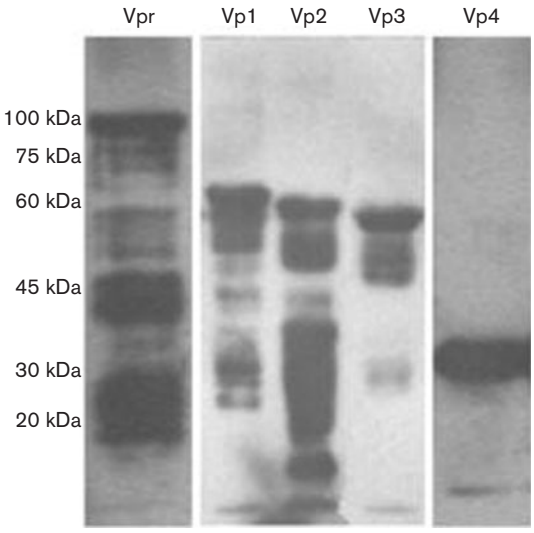

Fig. 4. *In vitro* auto-processing analysis of the full-length Vpr and deletion mutants Vp1–4. All the proteins were purified from inclusion bodies using Tris/HCl buffer, pH 8.0, containing 8 M urea. Contaminating bands were removed by using DEAE-Sephadex and Superdex 200/300 columns. Urea was finally removed by stepwise dialysis against Tris/HCl buffer, pH 8.0, with decreasing concentrations of urea (8 M to 0 M). Finally, refolded proteins were analysed by Western blotting using anti-Vpr antibody. The Vp1 mutant showed maximum auto-processing among the deletion mutants; the Vp4 mutant, which lacks the peptidaseS_8 domain, showed no auto-processing behaviour.

(Fig. 4). Furthermore, we found no auto-processing associated with the C-terminal mutant, Vp4 (Fig. 4). This was expected, as this mutant has no proteolytic activity. Previous observations by different research groups predicted both C-terminal and N-terminal auto-processing in Vpr (Park *et al.*, 2002; Sloma *et al.*, 1991). In the present study, we also found progressive auto-processing in Vpr from *B. cereus* DCUW.

DISCUSSION

The environment into which secretory proteins of Gram-positive bacteria emerge from the secretory translocase on the *trans* side of the cytoplasmic membrane plays an important role in their ultimate fate. Cleavable propeptides of various lengths and at different locations in the precursors of some secreted proteins represent significant folding factors. They are predominantly found at the N-terminal end of the primary translation product, between a typical signal peptide (prepeptide) and the N-terminus of the mature protein (Palacin *et al.*, 2002). Propeptides do not have a role in protein translocation across the membrane, which is initiated by the signal peptide, but they are essential in the post-translocational folding process to achieve the active and stable form of the secretory protein. For studies on the function of propeptides, bacterial serine proteases have served as the most important model proteins. In general, the propeptides of serine proteases serve two functions. Firstly, they are necessary for the rapid folding of these enzymes through a molten globule-like intermediate state. Secondly, under native conditions, propeptides behave as 'intramolecular chaperones' or foldases. However, they are also fully active in promoting folding when synthesized as separate molecular entities (Shinde & Inouye, 2000).

Recent evidence from different research groups has revealed that auto-processing of some high-molecular-mass serine proteases is involved in the processing of other propeptides (Corvey *et al.*, 2003; Okuda *et al.*, 2004) and thereby controls growth, metabolism and cell–cell communication in the extracellular medium. Minor extracellular protease Vpr has been implicated recently as a processor protease in the *Bacillus* system and it has been shown to process other secretory pro-proteins and propeptides [lantibiotic, subtilisin (Corvey *et al.*, 2003) and signal proteins (CSF; Lanigan-Gerdes *et al.*, 2007)]. Vpr has also been implicated in biofilm formation and quorum sensing (Lanigan-Gerdes *et al.*, 2007).

Sloma *et al.* (1991) identified and characterized Vpr from *B. subtilis* for the first time. The protein was found to show some unusual behaviours such as large protease complex formation. It undergoes C-terminal processing and results in multiple proteolytically active fragments. Recent evidence from B. C. Park's lab (Kho *et al.*, 2005) supports the previous observations on Vpr and they have reported Vpr as a fibrinolytic enzyme. In the present work have cloned and characterized Vpr from *B. cereus* DCUW, and screened the minimal Vpr domain that is essential for enzymic activity. According to previous observations, the long C-terminal domain in Vpr undergoes extensive auto-processing; to further confirm this, we made four C-terminal deletion mutants and showed experimentally that not only C-terminal processing but also N-terminal auto-processing is evident in Vpr. We found that the C-terminal deletion mutants were not capable of forming large protein complexes, which probably indicates that the long C-terminal protease-associated domain is involved in large molecular assembly formation. Further studies are necessary to understand the detailed mechanism of action of Vpr in the extracellular milieu. During the last decade, a considerable amount of information has been accumulating about proteins that have little or no ordered secondary structure under physiological conditions in the absence of their physiological partner/ligand. These proteins are referred to as 'natively unfolded' or 'intrinsically disordered' (Dunker *et al.*, 2001; Dunker & Obradovic, 2001). Our sequence analysis revealed that the long C-terminal domain in Vpr has nine intrinsically disordered regions. Secondary structural predictions showed that Vp4 (the C-terminal domain) contained 32.24 % extended strands, which explains the disorderedness in the C terminus. We

have validated the bioinformatics-based predictions experimentally. We found that the C-terminal domain in Vpr is not only unique in evolutionary pathways, but it is the domain that dictates the specific substrate recognition and also facilitates selecting partners in the extracellular milieu. This appears to be the first report on the sequence-based structure–function relationship of minor extracellular protease Vpr from the *B. cereus* group of organisms.

ACKNOWLEDGEMENTS

This study was supported by a grant from the CAS Programme of the Department of Biochemistry, University Grants Commission, India. Support for DNA sequencing was obtained from the NCMP grant to University of Calcutta, India. A. G. thanks the Guha Research Endowment Fund and the Department of Biotechnology (DBT) for financial assistance.

REFERENCES

Ageitos, J. M., Vallejo, J. A., Sestelo, A. B. F., Poza, M. & Villa, T. G. (2007). Purification and characterization of a milk-clotting protease from *Bacillus licheniformis* strain USC13. *J Appl Microbiol* **103**, 2205–2213.

Anson, M. L. (1938). The estimation of pepsin, trypsin, papain and cathepsin with hemoglobin. *J Gen Physiol* **22**, 79–89.

Corvey, C., Stein, T., Düsterhus, S., Karas, M. & Entian, K. D. (2003). Activation of subtilin precursors by *Bacillus subtilis* extracellular serine proteases subtilisin (AprE), WprA, and Vpr. *Biochem Biophys Res Commun* **304**, 48–54.

Dalbey, R. E., Lively, M. O., Bron, S. & van Dijl, J. M. (1997). The chemistry and enzymology of the type I signal peptidases. *Protein Sci* **6**, 1129–1138.

Dunker, A. K. & Obradovic, Z. (2001). The protein trinity – linking function and disorder. *Nat Biotechnol* **19**, 805–806.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, C. M., Ratliff, C. M. & other authors (2001). Intrinsically disordered protein. *J Mol Graph Model* **19**, 26–59.

Garnier, J., Gibrat, J. F. & Robson, B. (1996). GOR method for predicting protein secondary structure from amino acid sequence. *Methods Enzymol* **266**, 540–553.

Ghosh, A., Maity, B., Chakrabarti, K. & Chattopadhyay, D. (2007). Bacterial diversity of East Calcutta wet land area: possible identification of potential bacterial population for different biotechnological uses. *Microb Ecol* **54**, 452–459.

Ghosh, A., Chakrabarti, K. & Chattopadhyay, D. (2008). Degradation of raw feather by a novel high molecular weight extracellular protease from newly isolated *Bacillus cereus* DCUW. *J Ind Microbiol Biotechnol* **35**, 825–834.

Heussen, C. & Dowdle, E. B. (1980). Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. *Anal Biochem* **102**, 196–202.

Inouye, M. (1991). Intramolecular chaperone: the role of the propeptide in protein folding. *Enzyme* **45**, 314–321.

Itoi, Y., Horinaka, M., Tsujimoto, Y., Matsui, H. & Watanabe, K. (2006). Characteristic features in the structure and collagen-binding ability of a thermophilic collagenolytic protease from thermophile *Geobacillus collagenovorans* MO-1. *J Bacteriol* **188**, 6572–6579.

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapatral, V., Bhattacharyya, A., Reznik, G., Mikhailova, N. & other authors (2003). Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature* **423**, 87–91.

Khan, A. R. & James, M. N. G. (1998). Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci* **7**, 815–836.

Kho, C. W., Park, S. G., Cho, S., Lee, D. H., Myung, P. K. & Park, B. C. (2005). Confirmation of Vpr as a fibrinolytic enzyme present in extracellular proteins of *Bacillus subtilis*. *Protein Expr Purif* **39**, 1–7.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lanigan-Gerdes, S., Dooley, A. N., Faull, K. F. & Lazazzera, B. A. (2007). Identification of subtilisin, Epr and Vpr as enzymes that produce CSF, an extracellular signalling peptide of *Bacillus subtilis*. *Mol Microbiol* **65**, 1321–1333.

Lantz, M. S. & Ciborowski, P. (1994). Zymographic techniques for detection and characterization of microbial proteases. *Methods Enzymol* **235**, 563–594.

Letunic, I., Copley, R. R., Pils, B., Pinkert, S., Schultz, J. & Bork, P. (2006). SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res* **34**, D257–D260.

Ohta, Y. & Inouye, M. (1990). Pro-subtilisin E: purification and characterization of its auto-processing to active subtilisin E in vitro. *Mol Microbiol* **4**, 295–304.

Okuda, M., Sumitomo, N., Takimura, Y., Ogawa, A., Saeki, K., Kawai, S., Kobayashi, T. & Ito, S. (2004). A new subtilisin family: nucleotide and deduced amino acid sequences of new high-molecular-mass alkaline proteases from *Bacillus* spp. *Extremophiles* **8**, 229–235.

Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C. V. & Falquet, L. (2004). MyHits: a new interactive resource for protein annotation and domain identification. *Nucleic Acids Res* **32**, W332–W335.

Palacin, A., Parro, V., Geukens, N., Anne, J. & Mellado, R. P. (2002). SipY is the *Streptomyces lividans* type I signal peptidase exerting a major effect on protein secretion. *J Bacteriol* **184**, 4875–4880.

Park, S. G., Kho, C. W., Cho, S., Lee, D. H., Kim, S. H. & Park, B. C. (2002). A functional proteomic analysis of secreted fibrinolytic enzymes from *Bacillus subtilis* 168 using a combined method of two-dimensional gel electrophoresis and zymography. *Proteomics* **2**, 206–211.

Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* **57**, 50–108.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Res* **33**, W116–W120.

Sarvas, M., Harwood, C. R., Bron, S. & van Dijl, J. M. (2004). Post-translocational folding of secretory proteins in Gram-positive bacteria. *Biochim Biophys Acta* **1694**, 311–327.

Schmitz, S. K. (1990). *An Introduction to Dynamic Light Scattering by Macromolecules*. New York: Academic Press.

Shinde, U. & Inouye, M. (2000). Intramolecular chaperones: polypeptide extensions that modulate protein folding. *Semin Cell Dev Biol* **11**, 35–44.

Sloma, A., Rufo, G. A., Jr, Theriault, K. A., Dwyer, M., Wilson, S. W. & Pero, J. (1991). Cloning and characterization of the gene for an additional extracellular serine protease of *Bacillus subtilis*. *J Bacteriol* **173**, 6889–6895.

Strausberg, S., Alexander, P., Wang, L., Schwarz, F. & Bryan, P. (1993). Catalysis of a protein folding reaction: thermodynamic and

kinetic analysis of subtilsin BPN interactions with its propeptide fragment. *Biochemistry* **32**, 8112–8119.

von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. *Eur J Biochem* **133**, 17–21.

von Heijne, G. V. (1985). Signal sequences. The limits of variation. *J Mol Biol* **184**, 99–105.

Edited by: W. J. Quax