Protein Expression and Purification 96 (2014) 39–47

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Domain isolation, expression, purification and proteolytic activity of the metalloprotease PrtV from *Vibrio cholerae*

Aaron Edwin [a,b], Christin Grundström [a,b], Sun N. Wai [b,c,d], Anders Öhman [e], Gunter Stier [a,*], A. Elisabeth Sauer-Eriksson [a,b,*]

[a] Department of Chemistry, Umeå University, SE-901 87 Umeå, Sweden
[b] Umeå Centre for Microbial Research (UCMR), Umeå University, SE-901 87 Umeå, Sweden
[c] Department of Molecular Biology, Umeå University, SE-901 87 Umeå, Sweden
[d] The Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, SE-901 87 Umeå, Sweden
[e] Department of Pharmacology and Clinical Neuroscience, Umeå University, SE-901 85 Umeå, Sweden

ARTICLE INFO

Article history:
Received 17 December 2013
Available online 31 January 2014

Keywords:
Vibrio cholerae
Metalloprotease
PrtV
Fusion tag
Expression screen
Escherichia coli

ABSTRACT

The metalloprotease PrtV from *Vibrio cholerae* serves an important function for the bacteria's ability to invade the mammalian host cell. The protein belongs to the family of M6 proteases, with a characteristic zinc ion in the catalytic active site. PrtV constitutes a 918 amino acids (102 kDa) multidomain pre-pro-protein that so far has only been expressed in *V. cholerae*. Structural studies require high amounts of soluble protein with high purity. Previous attempts for recombinant expression have been hampered by low expression and solubility of protein fragments. Here, we describe results from parallel cloning experiments in *Escherichia coli* where fusion tagged constructs of PrtV fragments were designed, and protein products tested for expression and solubility. Of more than 100 designed constructs, three produced protein products that expressed well. These include the N-terminal domain (residues 23–103), the PKD1 domain (residues 755–839), and a 25 kDa fragment (residues 581–839). The soluble fusion proteins were captured with Ni^{2+} affinity chromatography, and subsequently cleaved with tobacco etch virus protease. Purification protocols yielded ~10–15 mg of pure protein from 1 L of culture. Proper folding of the shorter domains was confirmed by heteronuclear NMR spectra recorded on ^{15}N-labeled samples. A modified protocol for the native purification of the secreted 81 kDa pro-protein of PrtV is provided. Proteolytic activity measurements suggest that the 37 kDa catalytic metalloprotease domain alone is sufficient for activity.

Introduction

Cholera is caused by the motile Gram-negative bacterium *Vibrio cholerae,* that spreads via the oral-fecal route and on infection releases several virulence factors [1]. These factors include proteases that attack the target cells by breaking down tissue barriers and cellular matrix components, thereby causing necrosis [2–4]. One of these proteases is the secreted metalloprotease PrtV that exhibits a very potent cytotoxic effect. Known substrates for PrtV include blood plasma clotting components like fibrinogen, fibronectin and plasminogen, which help in immobilizing foreign entities [5,6].

PrtV belongs to the M6 peptidase family, sharing 37% sequence identity with the immune inhibitor A (InhA)[1] from *Bacillus thuringiensis*. The inactive PrtV is natively expressed as a 102 kDa full-length pre-pro-protein. Besides the signal sequence, PrtV consists of four domains: the N-terminal domain, the catalytic active M6 domain, and the two polycystic kidney disease domains, PKD1 and PKD2 (Fig. 1).

The N-terminal domain is present in many bacterial proteins, however, a specific function for it has not been identified. The M6 domain constitutes the catalytic metalloprotease domain with the characteristic HexxHxxgxxD Zn^{2+}-binding motif [7]. PKD domains are found in a variety of eukaryotic and prokaryotic proteins consisting of relatively short domains of 80–90 amino acids and are usually found in the extracellular parts of proteins involved

* Corresponding authors. Present address: Biochemie-Zentrum der Universität Heidelberg, Im Neuenheimer Feld 328, 69120 Heidelberg, Germany. Tel.: +49 6221 544789 (G. Stier). Address: Department of Chemistry, Umeå University, SE-90187 Umeå, Sweden. Tel.: +46 90 7865923 (A.E. Sauer-Eriksson).

E-mail addresses: gunter.stier@bzh.uni-heidelberg.de (G. Stier), elisabeth.sauer-eriksson@chem.umu.se (A.E. Sauer-Eriksson).

[1] *Abbreviations used:* InhA, immune inhibitor A; LB, Luria–Bertani; TEV, tobacco etch virus; MWCO, molecular weight cut off; NMR, nuclear magnetic resonance.

http://dx.doi.org/10.1016/j.pep.2014.01.012
1046-5928/

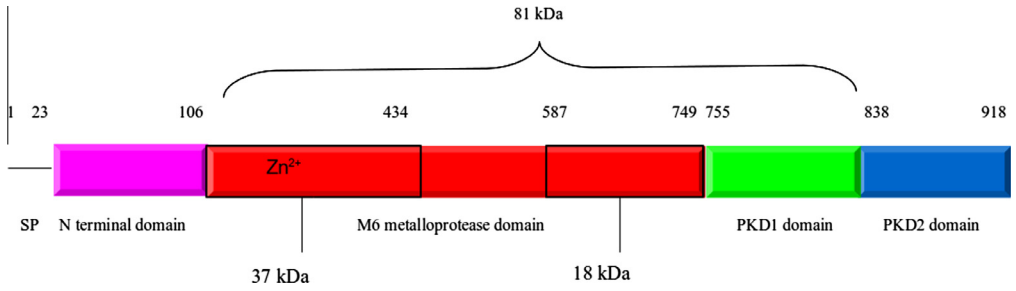

Fig. 1. Domain organization and maturation products of PrtV. A schematic representation of the four domains in the 102 kDa full-length PrtV.

in protein–protein or protein–carbohydrate interactions. Like the N-terminal domain, the function of the PKD domains in PrtV is not fully understood. It has been suggested that calcium binding at the PKD1 domain controls domain linker flexibility, and plays a regulatory role in the auto-proteolytic activity of the 81 kDa pro-protein [8].

The PrtV protease undergoes several N- and C-terminal modifications to form a catalytically active protease. The current model of the maturation process can be described: PrtV is secreted out of the bacteria as an 81 kDa pro-protein consisting of the M6 domain along with PKD1. This form of the protease is known to be stabilized by calcium ions [6]. Outside the bacterial cell, the 81 kDa form undergoes further cleavage to finally form the active M6 metalloprotease believed to comprise the 37 kDa and the 18 kDa domains, which together form the so called 55 kDa active complex [6].

We perform structural and functional studies of PrtV [5,6,8,9]. For this work, large quantities of pure and soluble protein are needed. In this study, we address the difficulties of protein domain expression of the PrtV protease by applying a divide and conquer approach. Screening of domain boundaries for individual domains and combinations of domains was carried out, and constructs obtained were further screened for solubility and expression levels. Recombinant PrtV fragments were obtained for three constructs including the N-terminal domain (residues 23–103), the PKD1 domain (residues 755–839), and a longer 25 kDa construct containing part of the M6 domain and PKD1 (residues 581–839) (Fig. 1). The previously described purification procedure for the 81 kDa pro-protein [6] was modified and improved to increase yield and purity. Interestingly, proteolytic activity measurements of purified domains suggest that the 37 kDa domain alone is sufficient for protein activity.

Methods

Construction of expression vectors

The pKVA232 plasmid with a cloned precursor of full-length PrtV (NP_232622.1) was used as a template for PCR amplification [6]. Various DNA sequences that code for individual or combinatorial PrtV domains were amplified using primers given in Table 1, which introduced 5′Nco1 and 3′Acc651 restriction sites for parallel cloning. The PCR products were cloned in parallel into the multiple cloning site of a pET24d plasmid coding for His_6-tagged protein and verified by DNA sequencing. Restriction fragments of constructs that expressed (Table 2, constructs marked yellow and green) were further sub-cloned into pET24d vectors coding for different N terminal His_6-carrier protein tags (Table 3) that would lead to PrtV constructs expressed as fusion proteins. Plasmids were produced in *Escherichia coli* DH5 alpha cells and stored at −80 °C.

Table 1
Primers used to amplify sequences of PrtV constructs.

Forward	
PrtV f	ATGAAAACGATCAAAAAAACGCTATTAGCTGCC
PrtV fN	GCTACCATGGCTAAAACGATCAAAAAAACGCTATTAGCTG
PrtV 23f	GCTACCATGGCTCAAACACCCATTGATTTAGGCGTGG
PrtV 34f	GCTACCATGGAGGATAAATTAATTGAAATGTTAGTCCG
PrtV 84f	GCTACCATGGAGCAGCGTGCGAAAATTCTTAAAGTG
PrtV 94f	GCTACCATGGATAAGCAAAAAGGCCCGCACAAG
PrtV 108f	GCTACCATGGATGTTGGTCAAAAGCGCACGGAC
PrtV 137f	GCTACCATGGAGCATACTGAGATGCTCTACGATCG
PrtV 430f	GGTACCATGGGCAAACGCGTTGAAGGCATTAAGCCTGCAGAA
PrtV 581f	GCTACCATGGCTCAAGGTTTCACCAAAAACGGTGGC
PrtV 591f	GCTACCATGGAAGCCAATCACTATTACTTGCTGC
PrtV 607f	GCTACCATGGACCAAGGCTTAGCCAATTTGAAACGC
PrtV 748f	GCTACCATGGGACGCTTGTCGAAAGTCACCACG
PrtV 755f	GCTACCATGGAAAATATCGCGCCTGTGGCT
PrtV 761f	GCTACCATGGCTCGCTTTGAACTGAAAGTCG
PrtV 769f	GCTACCATGGAGGGGCTCTCTGTGATGTC
PrtV 827f	GCTACCATGGATACTCATCAGCAAACCATTAAA
PrtV 838f	GCTACCATGGGTCCGAATGCGTTACCACAAGCCAG
Reverse	
PrtV 103r	GCTTGGTACCTTAACGCGCCTTGTGCGGGCCTTTTTGC
PrtV 115r	GCTTGGTACCTTAGTCCGTGCGCTTTTGACCAACATC
PrtV 137r	GCTTGGTACCTTACTCTTTCGTCAGGCGGTTATCATC
PrtV 425r	GCTTGGTACCTTATTTCACCATGTTCGGGCGTGAGTT
PrtV 434r	GCTTGGTACCTTATTCAACCCGTTTCATCGGCAGAG
PrtV 440r	GCTTGGTACCTTATTCTGCAGGCTTAATGCCTTCAAC
PrtV 452r	GCTTGGTACCTTAATCATCGCCACGGTTCGAGTAGAA
PrtV 573r	GCTTGGTACCTTATTCTGCGTTATCGATCAAGGTCTGATTG
PrtV 581r	GCTTGGTACCTTAACCTTGGAACGCAAAGCTGGATGTG
PrtV 591r	GCTTGGTACCTTATTCGTGGAAGCCACCGTTTTTGG
PrtV 607r	GCTTGGTACCTTAGTCAACGTCATTATGGCTGCGC
PrtV 751r	GCTTGGTACCTTATTTCGACAAGCGAACAACACCGTAC
PrtV 762r	GCTTGGTACCTTAGCGAGCCACAGGCGCGATATTTTCC
PrtV 768r	GCTTGGTACCTTACTCGACTTTCAGTTCAAAGCGAGC
PrtV 838r	GCTTGGTACCTTAATTCGGTGTGTCCACTTTAATGGTTTGC
PrtV rA	GCTTGGTACCTTACAGTTTGACTTTGATGGTGATGGTGGTG
PrtV rev	TTACAGTTTGACTTTGATGGTGATGGTGGTG

Expression screen of PrtV constructs

A single colony of freshly transformed *E. coli* Bl21 (DE3) pLysS strain (Novagen) containing one of the constructed plasmids was inoculated to 2 ml pre-culture of Luria–Bertani (LB) media containing 100 μg/ml kanamycin, 50 μg/ml chloramphenicol and 1% glucose, and grown overnight at 37 °C with shaking at 200 rpm. 0.5 ml from the pre-cultures were used to inoculate 9.5 ml of auto-induction LB media containing 100 μg/ml of kanamycin, 1 mM $MgSO_4$, 0.001% glucose, 0.02% lactose, 0.02% glycerol and 20 mM Tris pH 7.4 and grown at 37 °C with shaking. After 3 h cells were cooled to 25 °C and grown for 14–16 h with continued shaking. The cells were harvested by centrifugation at 4000*g* for 20 min and stored at −80 °C. Thawed cells were lysed by sonication and centrifuged at 18,000*g* for 30 min. Samples of the supernatant were analyzed by 15% SDS–PAGE to compare expression and solubility levels.

Table 2
Expression screen. The color code is as follows: red -- represents no expression, yellow -/+ represents low levels of expression, and green ++ represents high levels of expression.

	reverse	103	115	137	425	434	440	452	573	581	591	607	751	762	768	838	918
forward																	
23		++	-/+	-/+	--	--	--	--	--	--	--	--	--	--	--	--	--
34		--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
84					--	--	--	--	--	--	--	--	--	--	--	--	--
94					--	--	--	--	--	--	--	--	--	--	--	--	--
108					--	--	--	--	--	--	--	--	--	--	--	--	--
137					--	--	--	--	--	--	--	--	--	--	--	--	--
430									--	--	--	--	--	--	--	--	--
581													--	--	--	++	--
591													--	--	--	--	--
607													--	--	--	--	--
748																-/+	-/+
755																++	-/+
761																-/+	-/+
768																-/+	-/+
827																	--
838																	--

Table 3
Fusion tags used in the expression screen. The cloned and sequenced PCR product was shuttled into five different pET-system based vectors carrying different N-terminal His_6-tagged fusion partners as listed below. The His_6-tagged fusion partners were selected based on their reported properties to improve expression and solubility (for detailed information see: http://www.embl.de/pepcore/pepcore_services/strains_vectors/vectors/bacterial_expression_vectors).

Vector	Tag	Description	Reference
pETM-11	His_6–	His_6 tag	[20]
pETGB1_1a	His_6–GB1	His_6 + protein G binding domain	[21]
pETTrxA_1a	His_6–TrxA	His_6 + thioredoxin A	[22]
pETZZ_1a	His_6–ZZ	His_6 + protein A binding domain, double domain	[23]
pETMBP_1a	His_6–MBP	His_6 + maltose binding protein	[24,25]

Large scale expression and purification of His6–ZZ tagged PrtV 755-839 (PKD1 domain) and His6-tagged- and His6–TrxA-tagged PrtV 531-839 (25 kDa fragment)

Expression of the His_6–ZZ tagged PrtV PKD1domain and His_6-tagged- and His_6–TrxA-tagged PrtV 25 kDa fragment was done by upscaling the above described procedure to 4 L auto-induction media. All purification procedures were done on ice with pre-cooled solutions. Cell were resuspended in 10 ml per gram wet weight of lysis buffer consisting of buffer A (20 mM Tris pH 8.0 and 150 mM NaCl) supplemented with 10 mM imidazole, 1 mM pefabloc (Boehringer), 5 μg/ml lysozyme (Sigma–Aldrich) and 5 μg/ml DNAse 1 (Sigma–Aldrich). Resuspended cells were sonicated and insoluble material was separated by centrifugation at 40,000*g* for 40 min at 4 °C. The filtered supernatant (0.2 μm syringe filter, Sarstedt) was applied twice to 4 ml of Ni^{2+} NTA agarose (Qiagen) in econo columns (Biorad) that were pre-equilibrated with 5 column volumes of buffer A supplemented with 10 mM imidazole and 1 mM pefabloc (Boehringer). The column was washed with 5 column volumes of buffer A + 1 M NaCl followed by 5 column volumes of buffer A + 20 mM imidazole. The fusion protein was eluted with 5 column volumes of buffer A + 300 mM imidazole.

Tobacco etch virus (TEV) protease cleavage was done by adding a 1:100 M ratio of enzyme to the eluted fusion protein and allowing the mixture to dialyze against buffer A at 8 °C overnight. The dialyzed sample was passed through 4.5 ml column volume of Ni^{2+} NTA agarose (Qiagen) and the flow through containing the

target protein was collected. The protein solution was concentrated using 3 kDa molecular weight cut off (MWCO) centricon centrifugal filters (Millipore) and filtered through a 0.2 μm filter (Millipore) prior to loading onto a Superdex 75 16/60 (G.E. Healthcare) size exclusion chromatography column equilibrated with buffer A. The peak fractions were collected and analyzed by 15% SDS–PAGE. Both the PrtV 25 kDa fragment and the PrtV PKD1 domain protein samples were concentrated to 20 mg/ml. Total yield resulted in ~5 mg and ~15 mg from 1 L LB auto-induction media, respectively. The proteins were flash frozen in 40 μl aliquots in liquid nitrogen and stored at −80 °C.

The His_6-tagged- and His_6–TrxA-tagged PrtV 25 kDa fragment could not be cleaved with TEV protease. Thus the tag remains part of the purified protein.

Large scale expression and purification of His_6–TrxA tagged PrtV 23-103 (N-terminal Domain)

All steps for expression and purification of the His_6–TrxA tagged PrtV N domain construct were performed as described above up to the collection of the flow through from the second Ni^{2+} NTA agarose (Qiagen) column. The protein was dialyzed against buffer B (20 mM Tris pH 7.5 and 100 mM NaCl), filtered with 0.2 μm (Millipore) and loaded onto a MonoS 5/5 (G.E. Healthcare) column. The column was pre-washed with buffer B and the protein eluted by a gradient of 0–100% buffer B + 1 M NaCl. The protein fractions were analyzed by 15% SDS–PAGE and pooled together. The remaining purification protocol followed the same steps as described above. The PrtV N domain was concentrated to 20 mg/ml. From 1L media ~10 mg of protein was obtained. The protein was flash frozen in 40 μl aliquots in liquid nitrogen and stored at −80 °C.

In vivo expression and purification of PrtV 81 kDa pro-proteins

A fresh single colony of *V. cholerae* KAS202 strain [6] carrying the pKVA232 plasmid harboring a full length *PrtV* gene in an arabinose inducible system was inoculated to a starter culture of 50 ml low phosphate medium supplemented with 0.2% w/v casamino acids (Becton Dickinson and Franklin lakes, USA), 50 μg/ml L-tryptophan (Sigma–Aldrich), 20 mM NaCl, 1 mM $MgSO_4$, 5 mM $CaCl_2$ and 50 μg/ml carbenicillin [6]. Trace elements were also added at a final concentration of 25 μM $FeCl_3$, 10 μM $CaCl_2$, 10 μM $ZnSO_4$, 5 μM $MnCl_2$, 1 μM $CuCl_2$, 1 μM $CoCl_2$, 1 μM $NiCl_2$, 1 μM Na_2MoO_4, and 1 μM H_3BO_3 [10]. The culture was grown overnight at 30 °C with shaking at 200 rpm. The starter culture was used as inoculant to 10 L of the same media and allowed to grow at 30 °C with shaking until the cells reached an OD_{600} of 0.6 in the mid-exponential growth phase. Protein expression was induced by adding 0.01% w/v L-arabinose and expressed for 14–16 h with shaking at 30 °C. The cells were harvested by centrifugation at 8000*g* for 40 min. The collected supernatant, including the secreted 81 kDa pro-protein, was filtered through a 0.2 μm vacuum filter (Millipore), and loaded onto another 0.2 μm vacuum filter unit containing 10 ml DEAE bulk media (G.E. Healthcare) equilibrated with buffer C (20 mM Tris pH 8.0, 10 mM $CaCl_2$ and 20 mM NaCl). The media was then washed with 100 ml of buffer C and bound protein was eluted with 100 ml of buffer C + 500 mM NaCl. The eluted sample was subjected to sonication on ice and insoluble material was separated by centrifugation at 40,000*g* for 1 h. The collected supernatant was dialyzed against buffer C and loaded to a 5 ml Q sepharose column (G.E. Healthcare). The bound protein was washed with buffer C and eluted with buffer C + 500 mM NaCl.

Purification of the PrtV 37 kDa domain

Purification of the 81 kDa pro-protein was performed as described above and all buffers used hence forth in the purification of 37 kDa domain contained recommended concentration of the protease inhibitor cocktail complete-EDTA free (Roche). Purified 81 kDa PrtV pro-protein was dialyzed with buffer D (20 mM Tris pH 8.0 and 20 mM NaCl) at 25 °C for 16 h. The protein fragments were loaded onto a 5 ml Q sepharose column (G.E. Healthcare). The bound protein was washed with buffer D and eluted with buffer D + 500 mM NaCl. The eluted protein was concentrated using 3 kDa MWCO centricon centrifugal filters (Millipore) and a size exclusion chromatography was performed using a Superdex 200 16/60 column (G.E. Healthcare) equilibrated with buffer D + 100 mM NaCl. The fractions containing the 37 kDa domain were pooled and concentrated to 5 mg/ml. The purified protein was flash frozen in 40 μl aliquots in liquid nitrogen and stored at −80 °C.

Proteolytic activity assay of purified PrtV fragments

The purified N-terminal, 81, 37, 25 kDa (His_6-tagged), and PKD1 protein fragments were tested for their proteolytic activity. The calcium was removed from the protein by a quick buffer exchange over a PD 10 column (G.E. Healthcare) to buffer E (20 mM Tris pH 8.0 and 100 mM NaCl). Commercially available fibrinogen (Sigma–Aldrich) at a concentration of 1 mg/ml, in buffer F (50 mM HEPES pH 7.2 and 100 mM NaCl) was used to study the proteolytic activity of the respective PrtV domains. 100 nM of purified PrtV domains were added and incubated at 37 °C for 16 h. The boiled samples were subsequently analyzed by 15% SDS–PAGE.

Time scale proteolytic activity assay of the 81 kDa pro-protein

Calcium ions were stripped off the purified 81 kDa pro-protein by a quick buffer exchange over a PD 10 column (G.E. Healthcare) to buffer D. Immediately after the buffer exchange, 100 nM of the calcium-free protein fragment was added to 0.5 mg/ml fibrinogen in buffer F (time point = 0). The 81 kDa pro-protein was further incubated at 37 °C, and two additional samples of 100 nM each were taken after 2 and 16 h incubation, respectively, and mixed with 0.5 mg/ml fibrinogen in buffer F. All three fibrinogen samples with the added PrtV protease fragments were incubated at 37 °C for 1 h after which the samples were boiled and analyzed by 15% SDS–PAGE.

Isotope labeling and nuclear magnetic resonance (NMR) spectroscopy

Expression of His_6–TrxA PrtV N domain and His_6–ZZ PrtV PKD1 domain for NMR analysis were done as described earlier using 1 L of M9 minimal media supplemented with $^{15}NH_4Cl$ as the sole nitrogen source and trace elements as specified above. Labeled PrtV N domain and PrtV PKD1 domain were induced and purified as described above for the unlabeled protein. The proteins were exchanged into 50 mM potassium phosphate buffer, pH 7.4 and 500 mM NaCl. Unlabeled protein was also expressed using the same system in M9 minimal medium and purified as described.

NMR experiments were carried out on a Bruker DRX600 spectrometer, equipped with a 5 mm triple-resonance z-gradient cryo-probe. Fast-HSQC [11]. NMR experiments were recorded at 25 °C on ^{15}N-labeled 0.9 mM samples of PrtV fragments in a PBS buffer (pH 7.4) containing 10% D_2O. TOPSPIN, version 2.1 or 3.0 was used for spectrometer control, data processing and analysis (Bruker Biospin).

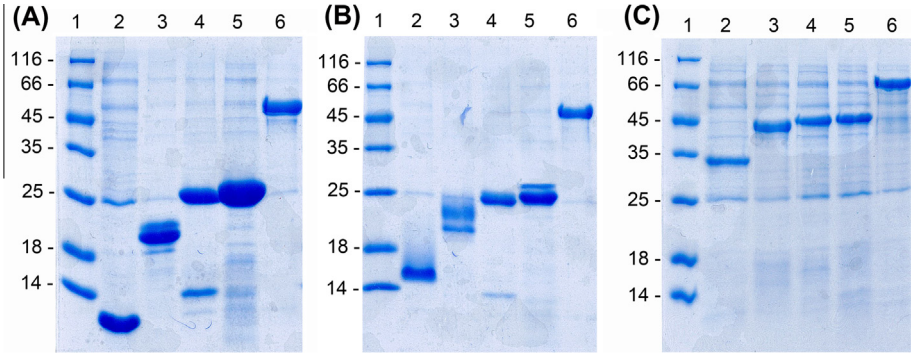

Fig. 2. Eluted proteins after nickel affinity chromatography. (A) The PrtV N terminal domain, (B) the PKD1 domain, and (C) the 25 kDa fragment with their fusion proteins followed by 15% SDS–PAGE. In each lane, 10 μl of the eluted samples were loaded as follows: Lane 1, protein marker; Lane 2, His_6-tag; Lane 3, His_6-GB1; Lane 4, His_6-TrxA; Lane 5, His_6-ZZ; Lane 6, His_6-MBP. All constructs show high levels of expression and solubility of the fusion proteins.

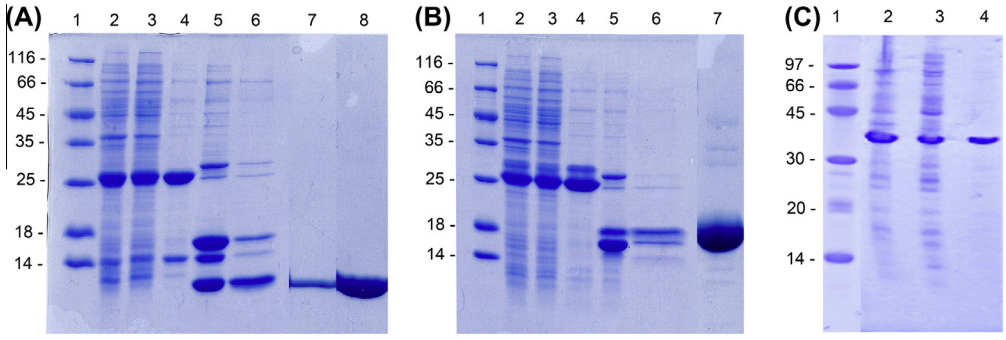

Fig. 3. (A) The purification profile of the His_6-TrxA tagged PrtV N terminal domain followed by 15% SDS–PAGE. Samples were loaded as follows: Lane 1, protein marker; Lane 2, total cell lysate; Lane 3, clarified lysate; Lane 4, elution from the first nickel affinity chromatography step; Lane 5, TEV digestion of the eluted protein; Lane 6, separated protein in the flow through after the second nickel affinity chromatography step; (Lanes 7 and 8 are taken from other gels) Lane 7, elution from the Mono S ion exchange chromatography; Lane 8, pooled and concentrated protein after size exclusion chromatography. (B) The purification profile of the His_6-ZZ tagged PrtV PKD1 domain analyzed with 15% SDS–PAGE. The loading scheme is similar to the first six lanes in (A); Lane 7 (from another gel), is the pooled and concentrated protein after size exclusion chromatography. (C) The purification profile of the His_6-TrxA-tagged PrtV 25 kDa fragment (total molecular weight equals 39.5 kDa including the 14.5 kDa His_6-TrxA tag) analyzed with 15% SDS–PAGE (Lane 1 is from another part of the gel). The loading scheme is similar to the first four lanes in (A).

Results

Expression, screening and domain identification

Various constructs of PrtV were produced corresponding to the previous reported domain borders and several inter-domain combinatorial variants (Table 2). Constructs that expressed were cloned into vectors expressing various carrier proteins as fusions (Table 3). The vector series was introduced to increase expression levels, proper folding and solubility of the target protein [12–14]. The fusion tag is linked to the target protein via a TEV protease recognition site that can be cleaved off with TEV protease during purification in order to produce untagged target proteins. Of more than 100 constructs, positive hits with high level of expression included three different constructs: the PrtV N domain (residues 23–103), the PrtV PKD1 domain (residues 755–838) and a PrtV 25 kDa fragment (residues 581–838 that includes the so called 18 kDa domain and the PKD1 domain) (Table 2, Fig. 1). These constructs were selected for further studies. Other constructs did not express or were insoluble or lost solubility after TEV cleavage.

Upscaled expression and purification

For upscaling expression and purification the His_6–TrxA tagged PrtV N domain, the His_6–ZZ tagged PrtV PKD1 domain and the His_6-tagged PrtV 25 kDa fragment fusion proteins were selected. The selection was based on the expression and solubility pattern of the proteins during the initial screen (Fig. 2). Previous studies suggest that the best choice of constructs for upscaling are the ones with the smallest fusion tag as this corresponds to the highest percentage of the target protein present in the total fusion protein [13,15]. The His_6 and His_6–GB1 fusion of PrtV N domain gave high yields, however a percentage of the expressed protein was also seen in the insoluble fraction. This was not seen with the His_6–TrxA fusion and consequently it was selected for upscaling (Fig. 2A). The same expression pattern was observed for the PrtV PKD1 domain and here the His_6–ZZ fusion construct was selected (Fig. 2B). In the case of the PrtV 25 kDa fragment construct, all fusions yielded similar amounts (Fig. 2C). The three selected protein fragments were purified with essentially the same purification protocol based on Ni^{2+} affinity chromatography, ion exchange chromatography, followed by size exclusion chromatography (Fig. 3A–C). The fusion partner of the PrtV N domain and the PrtV PKD1 domain could be removed with TEV protease and the target proteins were purified to homogeneity. However, the TEV protease failed to cleave the 25 kDa fragment, suggesting steric hindrance or a soluble aggregation of the protein [16].

NMR spectra were recorded to validate the structural integrity of the PrtV N domain and PrtV PKD1 domain. The [^{1}H-^{15}N]-HSQC spectra recorded at 25 °C were excellent where the number of resonances corresponds well to the number of residues of the respective domain (Fig. 4). While the moderate dispersion of resonances in the N-terminal domain suggests an α-helical protein the well dispersed resonances of the PKD1 domain is characteristic for a β-sheet protein.

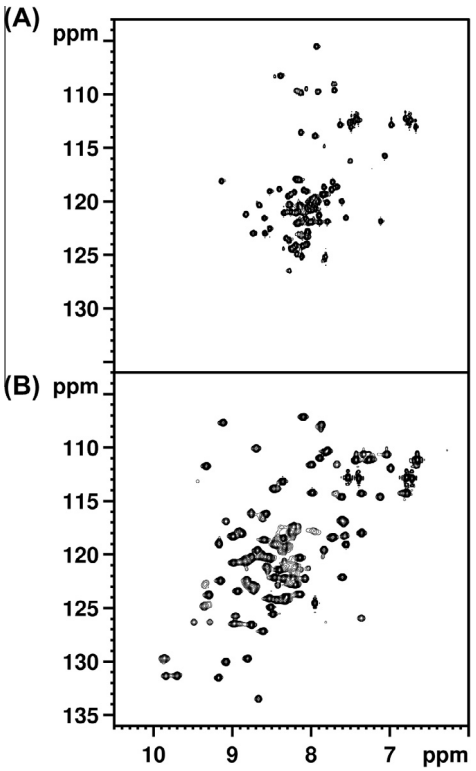

Fig. 4. [^{1}H-^{15}N]-HSQC NMR spectra of PrtV. The [^{1}H-^{15}N]-HSQC spectra of (A) a 0.9 mM ^{15}N-labeled N-terminal domain of PrtV, and (B) a 0.9 mM ^{15}N-labeled PKD1 domain of PrtV, recorded at 25°C.

Purification of the 81 kDa pro-protein from native source

The purification protocol of the secreted 81 kDa pro-protein from *V. cholerae* was modified to produce higher yields and protein of higher purity. The addition of the sonication treatment and high speed centrifugation at 40,000*g* were key factors contributing for increased purity. Using the earlier method, which did not include sonication and used a lower speed centrifugation at 20,000*g*, a highly viscous, mucous like phase, developed during the concentration of the protein. However, this was not seen with the protein samples purified with our modified method. The centrifugation step following the sonication treatment resulted in pellet formation which most probably comprised disrupted vesicles, membrane debris, misfolded and membrane bound proteins [17]. Size exclusion chromatography of the supernatant showed that the 81 kDa pro-protein purified in the form of high molecular weight oligomers (data not shown).

Autoproteolytic activity of the 81 kDa PrtV pro-protein

The autoproteolytic activity of the 81 kDa pro-protein was studied at various time points at 37 °C. The 81 kDa pro-protein is stable in the presence of calcium, and autoproteolysis can be initiated by dialysis against calcium free buffers [6]. In agreement with previous results our study shows that the 81 kDa pro-protein will immediately be auto-cleaved into two major fragments of 18 and 37 kDa, which after ~2 h are suggested to form the so called 55 kDa protein complex ([6], Fig. 5A, Lane 4). However, we now note that only the 37 kDa fragment appears to be stable during longer periods of time as the 18 kDa fragment undergoes further proteolysis and is almost undetectable after 16 h of incubation (Fig. 5A, Lane 5). The 37 kDa domain is stable and does not undergo further cleavage and could be purified with Q sepharose and size exclusion chromatography (Fig. 5B). Like the 81 kDa pro-protein, the 37 kDa domain was purified in a high molecular weight soluble form.

Proteolytic activity of PrtV fragments

The proteolytic activity of expressed and purified PrtV fragments was tested using commercially available fibrinogen, a known substrate of PrtV [6]. The 81 kDa pro-protein is inactive in the presence of calcium (Fig. 6A). By lowering the calcium concentration the proteolytic activity of the 81 kDa pro-protein was initiated. Interestingly, purified 37 kDa domains also turned out to be highly active, whereas, as expected, the PKD1 domain and the 25 kDa fragment did not show any sign of proteolytic activity (Fig. 6B). We carried out a time course to directly compare the catalytic activity of the 81 kDa pro-protein auto-cleavage products shown in Fig. 5A without any additional purification. In this study, samples were taken out at the various time points and mixed with

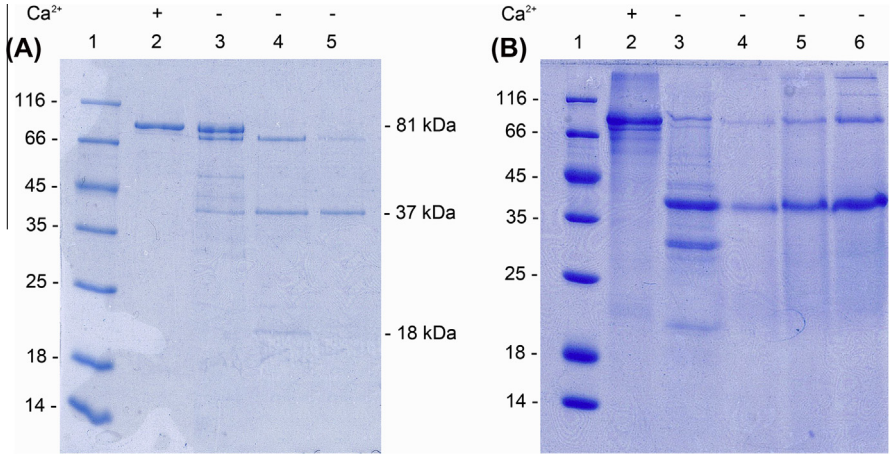

Fig. 5. (A) The autoproteolytic activity of the PrtV 81 kDa pro-protein in the absence of calcium followed by 15% SDS–PAGE. The loading scheme is as follows; Lane 1, protein marker; Lane 2, 81 kDa PrtV pro-protein; Lane 3, 81 kDa PrtV pro-protein after 5 min; Lane 4, 81 kDa PrtV pro-protein after 2 h; Lane 5, 81 kDa PrtV pro-protein after 16 h. (B) The purification of the PrtV 37 kDa fragment followed by 15% SDS–PAGE. The samples were loaded as follows: Lane 1, protein marker; Lane 2, the 81 kDa pro-protein; Lane 3, 81 kDa cleavage fragment loaded onto Q sepharose ion exchange chromatography; Lane 4, elution from the Q sepharose ion exchange chromatography; Lane 5, the protein sample loaded on size exclusion chromatography; lane 6, the concentrated elution peak from size exclusion chromatography. The molecular weights of the 37 kDa and 18 kDa domains were determined by mass spectroscopy [6]. Their positions on the SDS–PAGE are indicated.

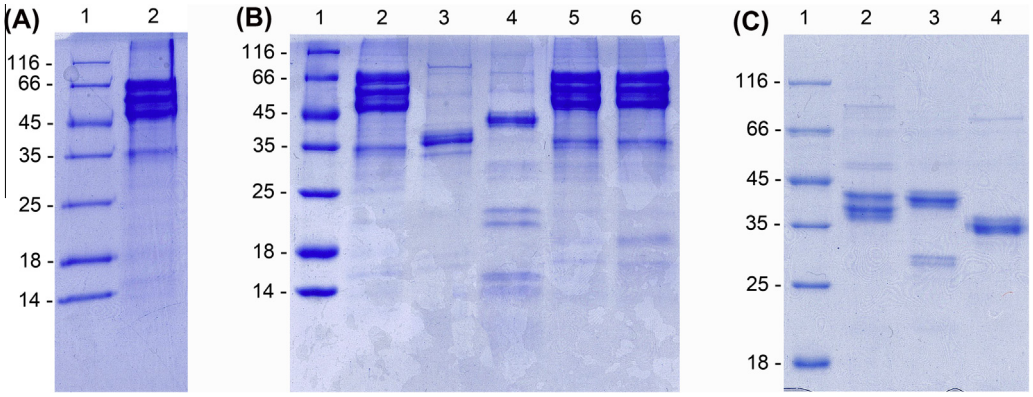

Fig. 6. Proteolytic activity of PrtV domains using fibrinogen as substrate. (A) Lane 1, protein marker; Lane 2, fibrinogen + 100 nM 81 kDa PrtV pro-protein + 10 mM $CaCl_2$. (B) The activity of PrtV fragments on fibrinogen followed by 15% SDS–PAGE. 10 μg of sample was loaded in each well as follows: Lane 1, protein marker; Lane 2, fibrinogen control; Lane 3, fibrinogen + 100 nM 81 kDa PrtV pro-protein; Lane 4, fibrinogen + 100 nM 37 kDa PrtV domain; Lane 5, fibrinogen + 100 nM 25 kDa PrtV protein; Lane 6, fibrinogen + 100 nM PrtV PKD1 domain. (C) 100 nM of the sample was added to 1 mg/ml fibrinogen, incubated at 30 °C for 1 h, and loaded on the gel as follows: Lane 1, protein marker; Lane 2, fibrinogen + 81 kDa pro-protein immediately after removal of calcium; Lane 3, fibrinogen + 81 kDa pro-protein incubated 2 h at 37 °C after removal of calcium; Lane 4, fibrinogen + 81 kDa pro-protein incubated for 16 h at 37 °C after removal of calcium.

fibrinogen. The results showed that the cleavage products after 16 h comprised predominantly the 37 kDa domain (Fig. 5A, Lane 5) and that this domain was equally, or perhaps even more, active in comparison with the activity of the fragments present after only 2 h of incubation (Fig. 5A, Lane 4 and 5; Fig. 6C, Lanes 3 and 4).

Discussion

The secreted metalloprotease PrtV is a potent virulence agent in the arsenal of *V. cholerae* causing instant cytotoxic effects during infection [6]. Several steps occur during the maturation process of the full length 102 kDa pre-pro-protein, undergoing the first cleavage steps in the bacterial cell before secretion as an 81 kDa intermediate pro-protein. By the removal of calcium this inactive pro-protein is auto-cleaved into smaller domains/fragments until it reaches its mature state as an active metalloprotease [6]. So far structural studies of natively produced and purified 81 kDa pro-protein have not been successful. We therefore set out to identify and purify individual domains of the protein, which requires that constructs of individual domains can be successfully created. Approximate domain borders determined by mass spectroscopy have been reported [6], however, optimized constructs that yield soluble proteins were not available.

In order to identify, overexpress and purify domains of PrtV, we employed several strategies that followed a divide and conquer approach. Our "domain phasing" and expression screening resulted in three target constructs: the PrtV N domain (residues 23–103), the PrtV PKD1 domain (residues 755–838) and the PrtV 25 kDa fragment (residues 581–838, which includes the so called 18 kDa fragment and the PKD1 domain). Interestingly, only slight differences in construct length resulted in these highly expressed and soluble constructs (Table 2). The PrtV N and PKD1 domains expressed well and could be purified as monomers in soluble form at high concentrations: ~20 mg of pure protein was obtained from 1 L media. The PrtV 25 kDa fragment could not be cleaved with TEV protease and could only be purified as high-molecular weight soluble aggregate.

In parallel with the expression screens we also optimized the purification protocol of the 81 kDa pro-protein secreted from its native source, i.e. the *V. cholerae* bacterium. Our aim here was to produce protein more suitable for crystallization purposes. However despite many attempts, crystals have so far not been obtained. The addition of the sonication treatment and high speed centrifugation at 40,000*g* to the previous purification protocol [6] removed a highly viscous, mucous like phase that otherwise developed during concentration of the protein. The centrifugation step after the sonication treatment resulted in pellet formation, which most probably included disrupted vesicles, membrane debris, misfolded and membrane bound proteins [17]. These impurities, now removed, are not visible on standard SDS–PAGE.

The autoproteolytic activity of the purified 81 kDa pro-protein was investigated. In the presence of calcium, the 81 kDa pro-protein is inactive, however, removal of calcium immediately starts an auto-proteolytic process and after 2 h at 37 °C, the 81 kDa pro-protein is degraded into the 37 and 18 kDa domains, i.e. the so called 55 kDa complex, in agreement with earlier reported findings [6] (Fig. 5A, Lanes 3 and 4). Interestingly, we found that with time the 18 kDa domain also will degrade leaving only the 37 kDa domain in solution (Fig. 5A, Lane 5). The 37 kDa domain (residues 106–434) is stable and could be purified with a combination of anion exchange and size exclusion chromatography (Fig. 5B).

We next asked if the 18 kDa domain is really necessary for protease activity. Therefore, the proteolytic activity of the 81 kDa pro-protein, as well as purified PrtV domains, was assayed using commercially available fibrinogen, a known substrate of PrtV [6]. The degradation of fibrinogen was monitored after 16 h of incubation of PrtV protein samples at 37 °C in a calcium-free solution (Fig. 6B). The results showed that purified 37 kDa domains had proteolytic activity towards the fibrinogen substrate. This suggests that the 18 kDa domain is not necessary for protease activity (Fig. 8). We also evaluated the proteolytic activity of the 37 kDa domain before purification (Fig. 6C). Noteworthy, the cleavage pattern of the fibrinogen substrate obtained by the 16 h auto-cleaved 81 kDa pro-protein and the purified 37 kDa domain are different from each other (compare Fig. 6B Lane 4 with Fig. 6C Lane 4). Even after incubation at 48 h the cleavage pattern of fibrinogen shown in Fig. 6B, lanes 3 and 4, does not change and fibrinogen is not cleaved further (data not shown). We cannot explain this difference. One could speculate that non-purified 37 kDa domains are associated with small molecular weight molecules/cleavage fragments that affect the activity of the protein and that these molecules are removed during purification. Size exclusion chromatography of purified 37 kDa domains shows that the protein is of high molecular weight. Thus observing catalytic activity suggests that this is not a case of soluble aggregation, rather that the fragments form high molecular weight multimers.

Based on auto-proteolysis and activity assays we propose a new model for PrtV maturation (Fig. 7). The 37 kDa active domain

```
         ---------E EEEEEEEE-- ---------- ---------- --HHHHHHHH
106      ALDVGQKRTD KVLALLIDFP DLPWDDNRLT KEHTEMLYDR YEPSHYQDLL

         ---------- ----HHHHHH HHHH---EEE EEEEEEEEEE ----------
         FSDKGYTGPN GENFISMRQY YESESGNSYS VSGQAAGWYR ASKNAAYYGG

         ---------- HHHHHHHHHH HHHH------ ---------- ----------
206      NSPGTNNDMN ARELVREALD QLARDPNINL ADYDIEDRYD YNGNGNFREP

         -- EEEEEEE EEE------- ---------- ---------- ----------
         DGVIDHLMIF HASVGEEAGG GVLGADAIWS HRFNLGRYHV LEGTKSNVPG

         ----EEEE-- ----------E EEEEHHHHH- ---------- ----------
306      RFNGQFAAFD YTIQPIDAAA GVCAHEYGHD LGLPDEYDTQ YTGTGEPVSY
                  D ++  P+  A   + AHE GH+ LG+  + + Q      E
         ADAM-8    D HSKNPVGVAC TM-AHEMGHN LGMDHDENVQ GCRCQERFEA

         EEE------- ---------- --HHHHHHHH H--------- EEEE------
         WSIMSSGSWA GKIGGTQPTA FSSWAKQFLQ NSIGGRWINH EQLSINELEA
            + +GS     IG + P   FS  ++ +L+ +
         GRCIMAGS-- --IGSSFPRM FSDCSQAYLE S

         ---------E E-----EEEE EE-------
406      KPRVVTLFQT TDNSRPNMVK VTLPMKRVE   434
```

Fig. 7. Calcium-dependent auto-proteolytic cleavage of the 81 kDa pro-protein results in formation of the 37 kDa active domain. The lightning bolt symbolizes the calcium-dependent cleavage between residues Leu749 and Ser750 [6].

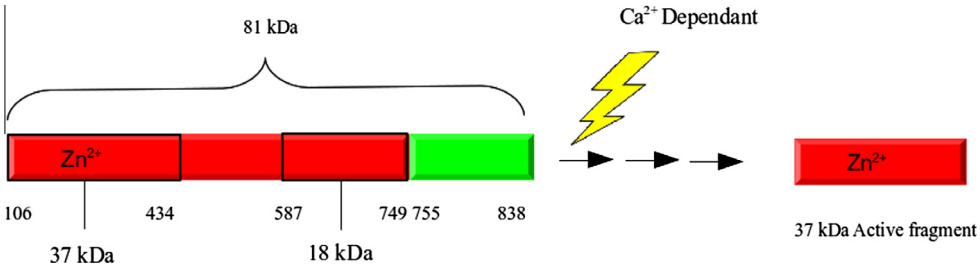

Fig. 8. Sequence alignment (Blast, [26] of the 37 kDa fragment of PrtV from *V. cholerae* and the human ADAM-8 domain. The sequence identity is 28% for 67 amino acids. Secondary structural elements from the ADAM-8 structure (pdb code 4dd8 [18]) are shown as red bars. The secondary structure elements predicted for PrtV with jpred3 [27] are shown in blue (E, β-strands) and red (H, α-helices). The conserved HexxHxxgxxD motif is highlighted in yellow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

includes residues 106–434 and constitutes the catalytic metalloprotease domain with the characteristic HexxHxxgxxD Zn^{2+}-binding motif [6,8]. A search at the protein data bank shows only 2 structures with low level of sequence identity to the 37 kDa domain. These include the human ADAM-8 metalloproteinase (pdb code 4dd8, [18]) and the zinc endoprotease from *Streptomyces caespitosus* (pdb code 1c7k, [7]). For the latter structure only 15 residues comprising the Zn^{2+}-binding motif are conserved. Sequence alignment with the ADAM-8 protein resulted in 28% identity over 67 aligned residues over the active site. It is therefore plausible that PrtV has a structure similar to ADAM-8 for residues Asp315–Asn386 (Fig. 8).

The N-terminal and PKD1 domains were expressed and purified to homogeneity. However, from the four constructs made of the PKD1 domain, only one including residues 755–838 could be expressed at high concentrations (Table 2). We have solved the crystal structure of the PKD1 domain at 1.1 Å resolution [8]. The domain forms a β-sandwich fold typical for PKD domains. A calcium binding site was identified in the PKD1 domain positioned close to its N-terminus. It was suggested that this calcium-binding site plays an important role in the regulation of the calcium-dependent auto-proteolysis mechanism [8]. Analysis of the domain boundaries as observed from the crystal structure of the PKD1 domain showed that any construct shorter than 757–838 would cut the first β-strands comprising the scaffold of the domain [8]. This would explain the lower protein expression levels and reduced protein stability of the constructs in the vicinity of this domain border (Table 2). The results support our approach to construct multiple variants of the same domain. Our strategy here also correlates with similar approaches from other labs that had constructed different domain lengths in order to achieve expression and solubility [13,14,19]. The same is true for the PrtV 25 kDa fragment where only one construct including residues 581–838 could be expressed (Table 2). This approach was however not fruitful in

the case of the 37 kDa domain from the M6 domain, which contains the catalytic active site. Still we have not been able to produce any 37 kDa domain recombinantly. Although *E. coli* is a Gram negative bacterium like *V. cholerae*, it may lack several key factors like chaperons that are necessary for protein folding and stability, and natively present in *V. cholerae*, a possibility that we will address in future experiments.

Acknowledgments

We thank Drs. Karina Persson, Uwe H. Sauer, and Tobias Hainzl for valuable ideas and discussions. The work was performed at the Umeå Centre for Microbial Research (UCMR) and was further supported by the Kempe Foundation and the Swedish Research Council; project K2013-67X-13001-15-3.

References

[1] J.B. Harris, R.C. LaRocque, F. Qadri, E.T. Ryan, S.B. Calderwood, Cholera, Lancet 379 (2012) 2466–2476.
[2] S. Miyoshi, S. Shinoda, Microbial metalloproteases and pathogenesis, Microbes Infect. 2 (2000) 91–98.
[3] C.C. Hase, R.A. Finkelstein, Bacterial extracellular zinc-containing metalloproteases, Microbiol. Rev. 57 (1993) 823–837.
[4] S. Shinoda, S. Miyoshi, Proteases produced by vibrios, Biocontrol Sci. 16 (2011) 1–11.
[5] K. Vaitkevicius, B. Lindmark, G. Ou, T. Song, C. Toma, M. Iwanaga, J. Zhu, A. Andersson, M.L. Hammarstrom, S. Tuck, S.N. Wai, A *Vibrio cholerae* protease needed for killing of Caenorhabditis elegans has a role in protection from natural predator grazing, Proc Natl Acad Sci U S A 103 (2006) 9280–9285.
[6] K. Vaitkevicius, P.K. Rompikuntal, B. Lindmark, R. Vaitkevicius, T. Song, S.N. Wai, The metalloprotease PrtV from *Vibrio cholerae*, FEBS J. 275 (2008) 3167–3177.
[7] G. Kurisu, T. Kinoshita, A. Sugimoto, A. Nagara, Y. Kai, N. Kasai, S. Harada, Structure of the zinc endoprotease from *Streptomyces caespitosus*, J. Biochem. 121 (1997) 304–308.
[8] A. Edwin, P. Rompikuntal, E. Bjorn, G. Stier, S.N. Wai, A.E. Sauer-Eriksson, Calcium binding by the PKD1 domain regulates interdomain flexibility in *Vibrio cholerae* metalloprotease PrtV, FEBS Open Bio. 3 (2013) 263–270.
[9] G. Ou, P.K. Rompikuntal, A. Bitar, B. Lindmark, K. Vaitkevicius, S.N. Wai, M.L. Hammarstrom, *Vibrio cholerae* cytolysin causes an inflammatory response in human intestinal epithelial cells that is modulated by the PrtV protease, PLoS ONE 4 (2009) e7806.
[10] F.W. Studier, Protein production by auto-induction in high density shaking cultures, Protein Expr. Purif. 41 (2005) 207–234.
[11] S. Mori, C. Abeygunawardana, M.O. Johnson, P.C. van Zijl, Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation, J. Magn. Reson. Ser. B 108 (1995) 94–98.
[12] W. Peti, R. Page, Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost, Protein Expr. Purif. 51 (2007) 1–10.
[13] J. Bogomolovas, B. Simon, M. Sattler, G. Stier, Screening of fusion partners for high yield expression and purification of bioactive viscotoxins, Protein Expr. Purif. 64 (2009) 16–23.
[14] N.S. Berrow, K. Bussow, B. Coutard, J. Diprose, M. Ekberg, G.E. Folkers, N. Levy, V. Lieu, R.J. Owens, Y. Peleg, C. Pinaglia, S. Quevillon-Cheruel, L. Salim, C. Scheich, R. Vincentelli, D. Busso, Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study, Acta Crystallogr. D Biol. Crystallogr. 62 (2006) 1218–1226.
[15] M. Hammarstrom, N. Hellgren, S. van Den Berg, H. Berglund, T. Hard, Rapid screening for improved solubility of small human proteins produced as fusion proteins in *Escherichia coli*, Protein Sci. 11 (2002) 313–321.
[16] D.S. Waugh, An overview of enzymatic reagents for the removal of affinity tags, Protein Expr. Purif. 80 (2011) 283–293.
[17] A. Kulp, M.J. Kuehn, Biological functions and biogenesis of secreted bacterial outer membrane vesicles, Annu. Rev. Microbiol. 64 (2010) 163–184.
[18] T. Hall, H.S. Shieh, J.E. Day, N. Caspers, J.E. Chrencik, J.M. Williams, L.E. Pegg, A.M. Pauley, A.F. Moon, J.M. Krahn, D.H. Fischer, J.R. Kiefer, A.G. Tomasselli, M.D. Zack, Structure of human ADAM-8 catalytic domain complexed with batimastat, Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun. 68 (2012) 616–621.
[19] G. Musco, G. Stier, C. Joseph, M.A. Castiglione Morelli, M. Nilges, T.J. Gibson, A. Pastore, Three-dimensional structure and stability of the KH domain: molecular insights into the fragile X syndrome, Cell 85 (1996) 237–245.
[20] V. Gaberc-Porekar, V. Menart, Perspectives of immobilized-metal affinity chromatography, J. Biochem. Biophys. Methods 49 (2001) 335–360.
[21] M. Hammarström, E.A.. Woesterenenk, N. Hellgren, T. Härd, H. Berglund, Effect of N-terminal solubility enhancing fusion proteins on yield of purified target protein, J. Struct. Funct. Genomics. 1 (2006) 1–14.
[22] P. Jurado, V. de Lorenzo, L.A. Fernandez, Thioredoxin fusions increase folding of single chain Fv antibodies in the cytoplasm of *Escherichia coli*: evidence that chaperone activity is the prime effect of thioredoxin, J. Mol. Biol. 357 (2006) 49–61.
[23] Y. Zhao, Y. Benita, M. Lok, B. Kuipers, P. van der Ley, W. Jiskoot, W.E. Hennink, D.J. Crommelin, R.S. Oosting, Multi-antigen immunization using IgG binding domain ZZ as carrier, Vaccine 23 (2005) 5082–5090.
[24] K.D. Pryor, B. Leiting, High-level expression of soluble protein in *Escherichia coli* using a His6-tag and maltose-binding-protein double-affinity fusion system, Protein Expr. Purif. 10 (1997) 309–319.
[25] S. Nallamsetty, D.S. Waugh, Solubility-enhancing proteins MBP and NusA play a passive role in the folding of their fusion partners, Protein Expr. Purif. 45 (2006) 175–182.
[26] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, J. Mol. Biol. 215 (1990) 403–410.
[27] C. Cole, J.D. Barber, G.J. Barton, The Jpred 3 secondary structure prediction server, Nucleic Acids Res. 36 (2008) W197–W201.