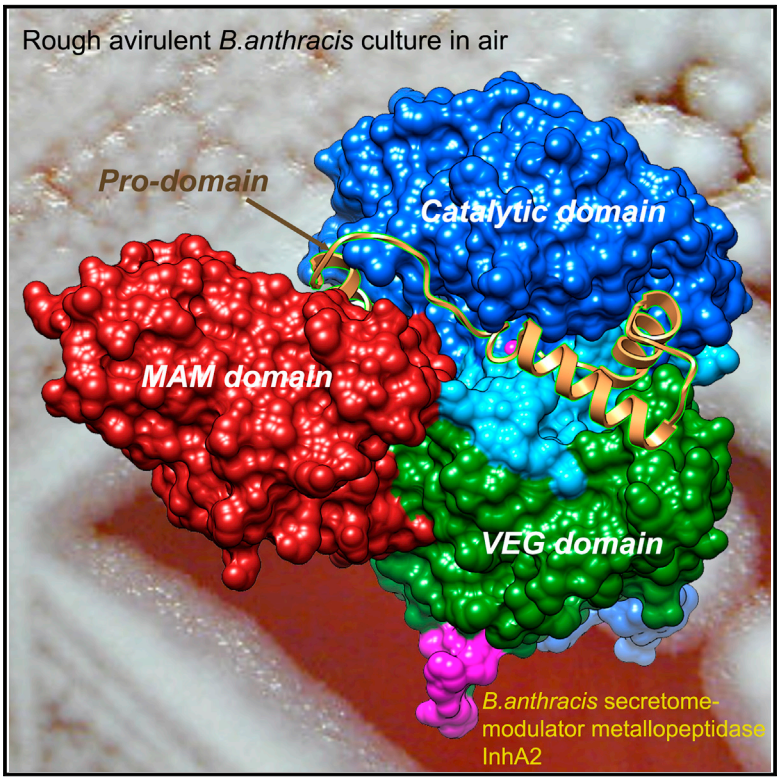
Article

Structure

Structural Basis for Latency and Function of Immune Inhibitor A Metallopeptidase, a Modulator of the *Bacillus anthracis* Secretome

Graphical Abstract



Highlights

- First structure of a *Bacillus anthracis* secretome-modulator InhA metallopeptidase
- Latency exerted via a unique N-terminal pro-domain blocking the active-site cleft
- Multi-domain protein with a VEG domain and a MAM domain found only in eukaryotes
- MAM domain grafted into the VEG domain is involved in folding and stability E-TOC

Authors

Joan L. Arolas, Theodoros Goulas, Andrei P. Pomerantsev, Stephen H. Leppla, F. Xavier Gomis-Rüth

Correspondence

xgrcri@ibmb.csic.es

In Brief

Immune inhibitor A(InhA)-type metallopeptidases (MPs) are secreted by *Bacillus cereus* group bacteria. *Bacillus anthracis* has *Ba*InhA1 and *Ba*InhA2, which degrade host tissue proteins. *Ba*InhA2 is a ~750-residue four-domain structure featuring a pro-peptide, a catalytic domain, a domain reminiscent of viral envelope glycoproteins, and a MAM domain, which is required for proper protein expression. Latency is exerted by the N-terminal segment of the pro-peptide, which binds the catalytic zinc.

Accession Numbers

4YU5
4YU6

Arolas et al., 2016, Structure *24*, 25–36
January 5, 2016
http://dx.doi.org/10.1016/j.str.2015.10.015

Structure
Article

Structural Basis for Latency and Function of Immune Inhibitor A Metallopeptidase, a Modulator of the *Bacillus anthracis* Secretome

Joan L. Arolas,[1,3] Theodoros Goulas,[1] Andrei P. Pomerantsev,[2] Stephen H. Leppla,[2] and F. Xavier Gomis-Rüth[1,*]
[1]Proteolysis Lab, Department of Structural Biology ("María de Maeztu" Unit of Excellence), Molecular Biology Institute of Barcelona, Spanish Research Council (CSIC), Barcelona Science Park, Helix Building, Baldiri Reixac, 15-21, 08028 Barcelona, Spain
[2]Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA
[3]Present address: Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna, Campus Vienna Biocenter 5, 1030 Vienna, Austria
*Correspondence: xgrcri@ibmb.csic.es
http://dx.doi.org/10.1016/j.str.2015.10.015

SUMMARY

Immune inhibitor A(InhA)-type metallopeptidases are potential virulence factors secreted by members of the *Bacillus cereus* group. Two paralogs from anthrax-causing *Bacillus anthracis* (*Ba*InhA1 and *Ba*InhA2) were shown to degrade host tissue proteins with broad substrate specificity. Analysis of their activation mechanism and the crystal structure of a zymogenic *Ba*InhA2 variant revealed a ~750-residue four-domain structure featuring a pro-peptide, a catalytic domain, a domain reminiscent of viral envelope glycoproteins, and a MAM domain grafted into the latter. This domain, previously found only in eukaryotes, is required for proper protein expression in *B. anthracis* and evinces certain flexibility. Latency is uniquely modulated by the N-terminal segment of the pro-peptide, which binds the catalytic zinc through its α-amino group and occupies the primed side of the active-site cleft. The present results further our understanding of the modus operandi of an anthrax secretome regulator.

INTRODUCTION

Bacillus anthracis is a facultative anaerobic Gram-positive spore-forming bacterium that belongs to the *Bacillus cereus* group (BCG) of Bacillales (Rasko et al., 2005). In 1876, Robert Koch identified it as the etiologic agent of anthrax (Koch, 1876), an ancient disease believed to have originated in Egypt and Mesopotamia around 1250 BC, which might have been among the ten plagues of Egypt (Sternbach, 2003). Anthrax is an acute, rapidly progressing infectious disease that affects humans and other animals (Baillie and Read, 2001), and infection occurs through the spores, which produce three clinical pictures depending on the entry pathway: cutaneous, inhalational, or gastrointestinal anthrax. The former leads to skin ulcers, which are generally easy to treat. In contrast, inhalational and gastrointestinal anthrax are invasive and systemic, and often fatal. Accordingly, novel therapeutic approaches targeting *B. anthracis* are of importance and currently under investigation (Artenstein and Opal, 2012).

In addition to the major virulence factors of *B. anthracis* identified, namely the plasmid-encoded anthrax toxin and the poly-D-glutamic acid capsule, other proteins contribute to virulence and disease, and may provide targets for therapy against anthrax (Artenstein and Opal, 2012). Among candidate virulence factors are immune inhibitor A (InhA) peptidases, which belong to the thuringilysin family within the metzincin clan of metallopeptidases (MPs) (Cerdà-Costa and Gomis-Rüth, 2014). They have also been ascribed to family M6 in the MEROPS proteolytic enzyme database (http://merops.sanger.ac.uk). Metzincins are characterized by a globular catalytic domain (CD) spanning ~130–270 residues, which consists of a structurally conserved N-terminal upper subdomain (NTS) and a lower, structurally more disparate C-terminal subdomain (CTS). The two subdomains are separated by the active-site cleft. NTSs span a (mostly) five-stranded β sheet, a backing helix, and an active-site helix, which contains an extended metal-binding motif, **H***E*XX**H**XX(G,N)XX(**H,D**). This motif encompasses three metal-binding residues (in bold) and the catalytic general base/acid glutamate (in italics) required for reaction. Some metzincins also include an additional adamalysin helix, first described in adamalysins/ADAMs (Cerdà-Costa and Gomis-Rüth, 2014). Downstream CTSs mainly share a conserved loop with a central methionine, the Met-turn, which provides a hydrophobic basis for the metal-binding site (Tallant et al., 2010), and a C-terminal helix. Among the 12 metzincin families structurally characterized are the toxilysins (Cerdà-Costa and Gomis-Rüth, 2014; Gomis-Rüth, 2013; Ng et al., 2013), which are part of the bacterial AB_5 toxins (Beddoe et al., 2010). Like the protective antigen and lethal factor tandem in *B. anthracis* anthrax toxin, AB_5 toxins from *Escherichia coli* (EcxAB) and *Citrobacter freundii* (CfxAB) consist of a subunit for host-cell invasion (pentameric subunit B) and an active MP that subverts intracellular functions (subunit A) (Beddoe et al., 2010).

The thuringilysins are one of the oldest known MP families, as the founding member, InhA1 from *Bacillus thuringiensis*, was identified in the mid-1970s as an immune inhibitor that interferes

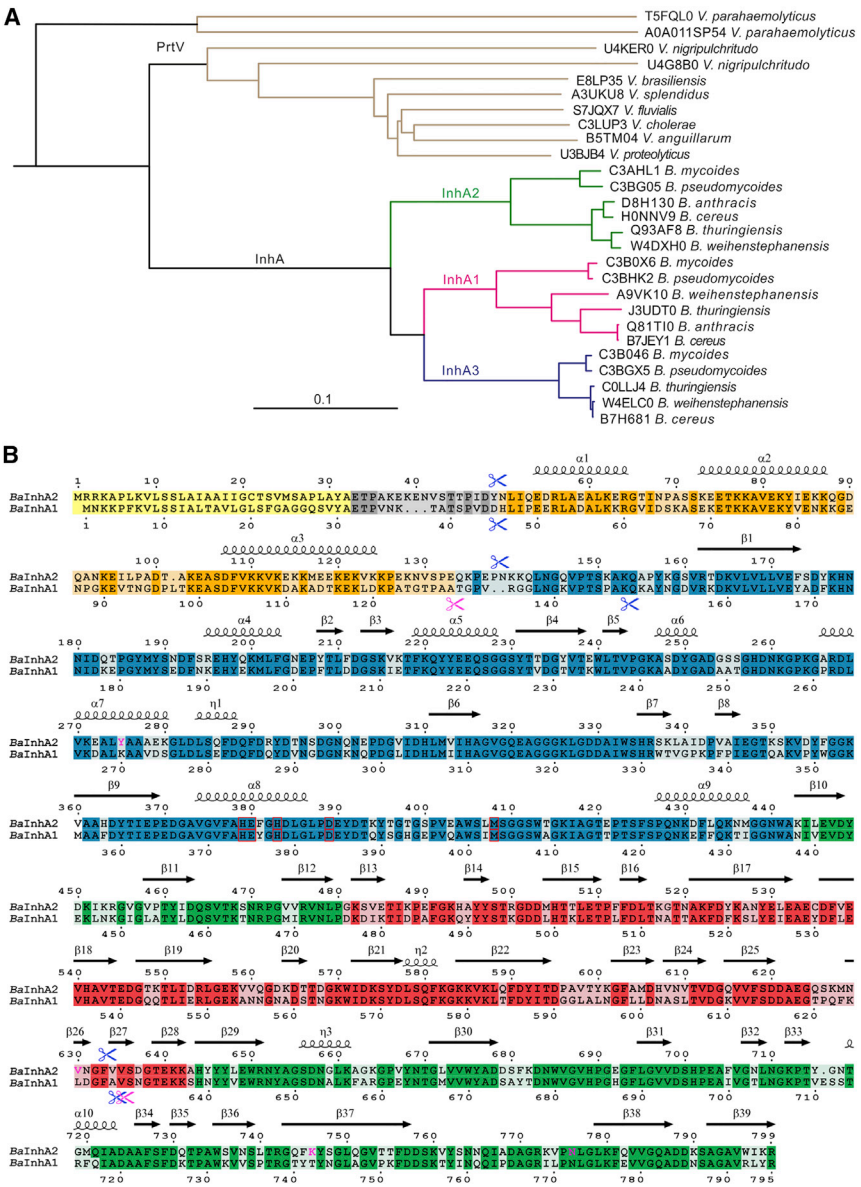


Figure 1. Phylogenetic Analysis and Sequence Alignment of Thuringilysins

(A) Phylogenetic tree of selected *Bacillus* (InhA) and *Vibrio* (PrtV) thuringilysins, each with its UniProt code. The bar represents 0.1 substitutions per site. The tree shows three major groups of InhA peptidases: A1 (magenta), A2 (green), and A3 (blue).

(B) Sequence alignment of *Ba*InhA1 (UniProt: Q81TI0) and *Ba*InhA2 (UniProt: D8H130). Signal peptides are in yellow, pro-peptides in orange, catalytic domains in blue, VEG domains in green, and MAM domains in red. Identical residues are shown over dark background. Autolytic cleavage sites are indicated by scissors (in blue, this work; in magenta, see Chung et al., 2006). Residues involved in zinc coordination and catalysis, as well as in the Met-turn, are boxed in red. Residues of *Ba*InhA2 cloned from *B. anthracis* strain Ames 33 (this work) that differ from those deposited (strain CI, UniProt: D8H130) are colored magenta. Regular secondary structure elements of the *Ba*InhA2 structure are shown.

See also Figure S3.

with the humoral defense system of saturniid pupae in the hemolymph by degrading antibacterial insect proteins (Edlund et al., 1976). Close homologs with >60% pairwise sequence identity, including InhA1, InhA2, and InhA3 paralogs, are found within BCG members (Charlton et al., 1999), and a possible explanation for this paralogy seems to be complementary expression patterns during the distinct stages of the *Bacillus* life cycle (Guillemet et al., 2010). Slightly more distant thuringilysin family members (Gomis-Rüth, 2003) include PrtV peptidases from several *Vibrio* species, including human pathogenic *Vibrio cholerae* (*Vc*PrtV; Vaitkevicius et al., 2008) and *Vibrio parahaemolyticus*, which show >40% sequence identity with *Bacillus* counterparts (Figure 1A).

Among thuringilysins are also InhA1 and InhA2 from *B. anthracis* (*Ba*InhA1 and *Ba*InhA2, respectively). These are synthesized as pre-pro-peptidases of 795 and 799 residues, respectively (69% sequence identity; Figure 1B), with signal peptides (SPs) for secretion of the zymogenic forms. Extracellular activation of the latter prevents damage to the bacterium. In general, the transcription levels of *inhA1* are 50–100 times higher than those of *inhA2*, which can be explained by the presence of a PlcR-binding site upstream of the promoter region of the latter gene (McKenzie et al., 2014). *Ba*InhA1 is highly abundant in supernatants of cultures grown under either aerobic or anaerobic conditions (Chitlaru et al., 2006), which can be found in soil environment or within mammalian host tissues during infection. *Ba*InhA1 modulates, directly or indirectly, the form and abundance of more than half of the secreted proteins of *B. anthracis* (Pflughoeft et al., 2014). In addition, it participates in nutrient acquisition by digesting host-serum proteins such as hemoglobin (Terwilliger et al., 2015). As to its function as a candidate virulence factor, *Ba*InhA1 contributes to anthrax pathology through direct degradation of host tissues: it cleaves various extracellular matrix proteins (Chung et al., 2006), regulates adhesion in human endothelial cells (Tonry et al., 2012), and increases the blood-brain barrier disruption associated with anthrax meningitis (Mukherjee et al., 2011). This MP also affects the coagulation cascade by degrading thrombin-activatable fibrinolysis inhibitor, von Willebrand factor, plasma ADAMTS-13, prothrombin, and factor X, among others (Chung et al., 2011).

To better understand the function of thuringilysins, we examined activation and proteolytic activity of *Ba*InhA1 and *Ba*InhA2 peptidases in vitro, and found that both were competent MPs of broad specificity against proteins and peptides. We further determined the crystal structure of the *Ba*InhA2 zymogen, which

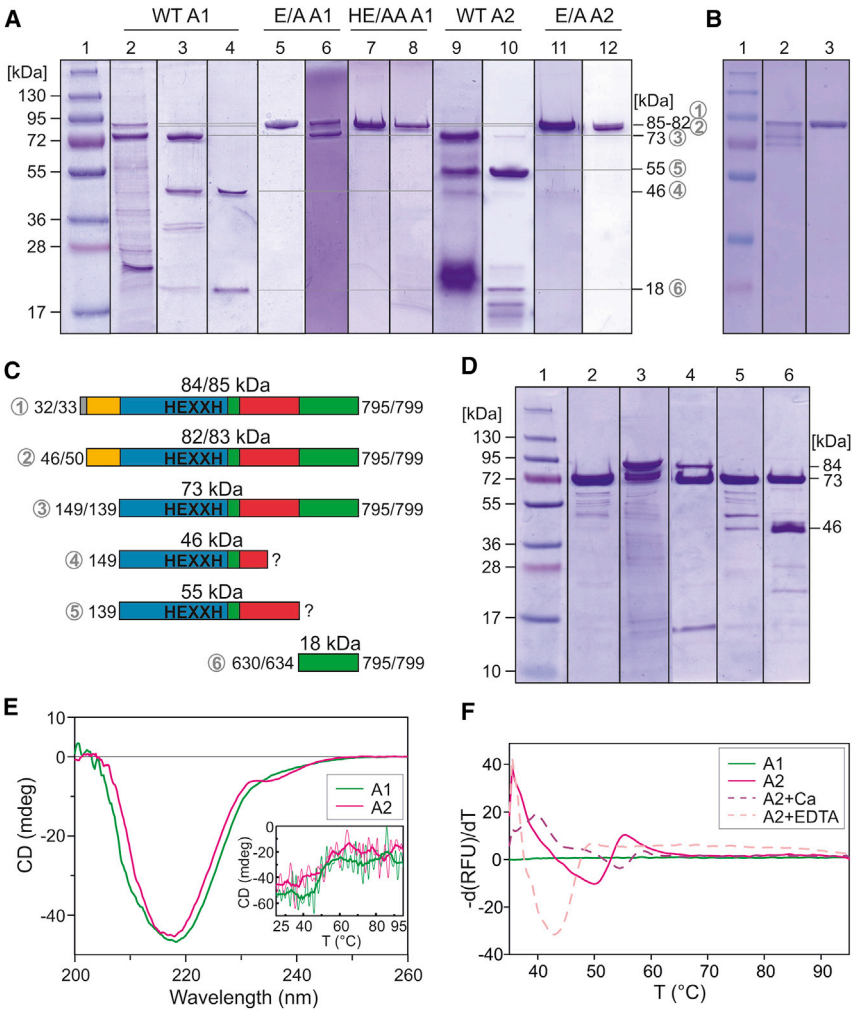


Figure 2. Activation and Stability of *Ba*InhA1 and *Ba*InhA2 Peptidases

(A) SDS-PAGE of WT, single mutant (E/A), and double (HE/AA) mutants of *Ba*InhA1 and *Ba*InhA2. Lane 1, molecular mass standard (mms). Lane 2, culture supernatant of WT *Ba*InhA1 (similar results for WT *Ba*InhA2). Lanes 3, 5, 7, 9, and 11, WT *Ba*InhA1, *Ba*InhA1-E/A, *Ba*InhA1-HE/AA, WT *Ba*InhA2, and *Ba*InhA2-E/A, respectively, after affinity purification. Lanes 4, 6, 8, 10, and 12, WT *Ba*InhA1, *Ba*InhA1-E/A, *Ba*InhA1-HE/AA, WT *Ba*InhA2, and *Ba*InhA2-E/A, respectively, after SEC and incubation either at room temperature (WT) or 4°C (mutants) for autolytic maturation.

(B) Autolytic processing of *Ba*InhA1-E/A and -HE/AA mutants. Lane 1, mms. Lane 2, *Ba*InhA1-E/A. Lane 3, *Ba*InhA1-HE/AA (both at 0.2 mg/ml) after SEC and incubation at 4°C.

(C) Correlation between bands observed in (A) and domain structure of *Ba*InhA1 and *Ba*InhA2 (same colors as in Figure 1B). The autolytic events are labeled with gray numbers as in (A). Sequences according to UniProt: Q81TI0/D8H130.

(D) Limited proteolysis of *Ba*InhA1-E/A. Lane 1, mms. Lanes 2–4, protein incubated at 1:100 peptidase/substrate weight ratio with trypsin, chymotrypsin, and thermolysin, respectively. Lanes 5 and 6, protein incubated with trypsin and thermolysin, respectively, at a 10:1 weight ratio.

(E and F) Circular dichroism spectra of mature WT *Ba*InhA1 and *Ba*InhA2 (lanes 4 and 10 in A, respectively) at 25°C (E). Unfolding transition curves showing temperature-dependent change in fluorescence (F) and in circular dichroism signal at 215 nm (inset in E) for both mature proteins.

See also Figure S1.

revealed four distinct domains and a novel mechanism of latency among MPs. Taken together, these data shed light on the latency, activation, and activity of the peptidase family and, thus, lay the foundation for specific drug design against anthrax.

RESULTS AND DISCUSSION

Recombinant Production and Autolytic Activation of *B. anthracis* InhA1/A2 Peptidases

Heterologous expression trials of several constructs in *E. coli* (see Supplemental Experimental Procedures) yielded insoluble protein under all conditions tested. We circumvented this problem by producing full-length *Ba*InhA1 (84 kDa; residues E_{32}–R_{795}; UniProt: Q81TI0; residues numbered in subscript) and *Ba*InhA2 (85 kDa; residues E^{33}–R^{799}; UniProt: D8H130; residues numbered in superscript) lacking the predicted endogenous SPs in a *Bacillus subtilis* strain deficient in eight proteases with a vector that provides a *Bacillus* SP for extracellular secretion instead. Recombinant wild-type (WT) *Ba*InhA1 and *Ba*InhA2 underwent fast stepwise processing (Figure 2A). This was strongly diminished in *Ba*InhA1 and *Ba*InhA2 active-site mutants, in which the catalytic glutamates were replaced by alanine, thus resulting in residual activity only (E/A mutants; Figure 2A); and was completely abolished in a double-point mutant affecting both the catalytic glutamate and a zinc-binding histidine of *Ba*InhA1, which was inactive (HE/AA mutant; Figure 2A). This indicated that processing was autolytic, as previously described for other recombinant bacterial MPs upon purification and concentration, such as karilysin from *Tannerella forsythia* (Cerdà-Costa et al., 2011), fragilysin-3 from *Bacteroides fragilis* (Goulas et al., 2011), and snapalysin from *Streptomyces coelicolor* (Trillo-Muyo et al., 2013).

Analysis of purified *Ba*InhA1 and *Ba*InhA2 by SDS-PAGE, combined with N-terminal sequencing and mass spectrometry (see Supplemental Experimental Procedures), revealed several species spanning between 18 and 84/85 kDa (*Ba*InhA1/*Ba*InhA2; Figures 2A and 2B). In particular, the latter species (circled 1 in Figures 2A and 2C), with disordered N-terminal segments as predicted by bioinformatics and indicated by small-angle X-ray scattering studies (SAXS) in solution (see below), were transformed over time into stable zymogens (circled 2 in Figures 2A and 2C) after cleavage before H_{46} and N^{50}, respectively. This was the form that crystallized for *Ba*InhA2 (see below). A second series of cleavages removed the zymogenic

pro-peptides (H_{46}–A_{129} and N^{50}–E^{133}, respectively; for reference, mature purified *Ba*InhA1 starts at T_{130}; Chung et al., 2006; see Figures 2A and 2C) plus some extra residues of the mature moiety. This led to the single-chain mature forms (circled 3 in Figures 2A and 2C), which were equivalent to purified natural *Ba*InhA1 and *Ba*InhA2, respectively. Finally, subsequent cleavages, most notably one before V_{630} and V^{634}, respectively (Figures 2A and 2C), resulted in the final mature forms, which consisted of two non-covalently associated fragments of 46/55 kDa (*Ba*InhA1/*Ba*InhA2) and 18 kDa (circled 4 and 5, plus 6 in Figures 2A and 2C). The two fragments that build up the final mature forms (46 + 18 kDa in *Ba*InhA1, consistent with purified protein from *B. anthracis* extracts, see Chung et al., 2006; and 55 + 18 kDa in *Ba*InhA2) separated in SDS-PAGE but not in size-exclusion chromatography (SEC), i.e. they strongly but non-covalently associated. The activation of *Ba*InhA1 and *Ba*InhA2, however, may follow a heterolytic mechanism in vivo triggered by other secreted bacterial peptidases or host peptidases at the site of infection. Indeed, limited proteolysis of *Ba*InhA1 E/A mutant with trypsin, chymotrypsin, and thermolysin likewise resulted in pro-peptide removal (Figure 2D). Importantly, the final mature forms obtained from *B. subtilis* evinced a well-folded conformation as assessed by circular dichroism, which showed similar spectra and denaturation for both *Ba*InhA1 and *Ba*InhA2 (Figure 2E). Their stability in front of denaturation was further examined by the thermofluor approach, which revealed no dye binding to the former and a temperature of mid-transition of 52°C for the latter (Figure 2F).

Proteolytic Activity and Inhibition of *B. anthracis* InhA1/A2

Mature recombinant *Ba*InhA1 and *Ba*InhA2 efficiently digested natural protein substrates including casein, fibrinogen, fibronectin, actin, and gelatin, but not BSA or α-lactalbumin, which is a highly stable multi-disulfide protein (Figure 3A). *Ba*InhA1 further digested β-lactoglobulin and, notably, the major structural component of the mucus, MUC2 mucin, which covers and protects the epithelium of the intestine from bacteria (Johansson et al., 2013). MUC2 mucin was cleaved with a profile similar to that of the human intestinal MP meprin β (compare Figure 3A, lower right and Figure 2 in Schutte et al., 2014). Thus, *Ba*InhA1 could have a role in the disruption of the dense gel that mucins form, permitting anthrax to overcome this gastrointestinal barrier and attack the intestinal epithelium during infection. *Ba*InhA1 further digested authentic triple-helical collagen at multiple sites like a bona fide collagenase, thus breaking down a structural protein that may be crucial to all three anthrax infection processes. *Ba*InhA1 was more active than *Ba*InhA2 against all protein substrates analyzed. This also held for fluorogenic casein and gelatin, which were cleaved by *Ba*InhA1 ~5-fold more efficiently than by *Ba*InhA2 (Figure 3C). WT *Ba*InhA1 and *Ba*InhA2 were also overproduced extracellularly by a *B. anthracis* strain deficient in six proteases (Pomerantsev et al., 2011), and they showed equivalent activity to that of the proteins expressed in *B. subtilis* on fluorogenic casein (Figures 2C, S1A, and S1D). As to peptidolytic activity, *Ba*InhA1 cleaved peptide Mca-R-P-K-P-V-E-Nva-W-R-K(dnp)-NH_2 with high efficiency and Mca-P-L-A-Nva-Dap(dnp)-A-R-NH_2 less efficiently, whereas *Ba*InhA2 only digested the former. The rest of peptides tested were not affected with the exception of Abz-V-K-F-Y-D-I-K(dnp), which was digested by *Ba*InhA1, thus suggesting broad specificity, in particular against peptides. In addition, both *B. anthracis* proteins displayed maximum proteolytic activity at pH 6.5 (Figure 3D), which slightly deviates from previous results with purified *Ba*InhA1 (pH 7–8; Chung et al., 2006). The MPs were inhibited only by general MP inhibitors but not by specific inhibitors of other MPs or peptidase classes (Table S2), and by the universal pan-peptidase inhibitor α_2-macroglobulin (Figure 3B).

Accordingly, recombinant *Ba*InhA1 and *Ba*InhA2 were competent MPs, which showed overall broad cleavage specificity against proteins and peptides, as also found during autolytic activation (see above) and in natural purified *Ba*InhA1 protein (Chung et al., 2006). This broad specificity is consistent with the function of *Ba*InhA1 in the *B. anthracis* secretome, where it digests not only endogenous toxins (protein antigen, lethal factor, and edema factor) and other peptidases (thermolysin Npr599), but also multiple chitinases and chitin-binding proteins which enable bacteria to thrive in the soil (Pflughoeft et al., 2014).

Overall Structure of *B. anthracis* InhA2 Zymogen

A number of constructs and mutants of *Ba*InhA1 and *Ba*InhA2 were assayed for crystallization, but only *Ba*InhA2 E/A mutant zymogen (hereafter pro-*Ba*InhA2-E/A) crystallized adequately for structure analysis. However, given the high sequence identity between the two paralogs (69%; see Figure 1B), the structure-derived findings hereafter may be valid for both. The structure was solved by a two-wavelength anomalous dispersion experiment with a selenomethionine (SeMet) derivative, and final models were obtained and refined against derivative data to 2.9 Å and native data to 2.6 Å resolution, respectively (see Table 1). Overall, structural information is available for two molecules (A and B) from the native crystals and two from the derivative (A and B). Given that molecule A of the native structure is the most complete of the four (see Supplemental Experimental Procedures), it is taken as a reference hereafter. This molecule spans the full segment N^{50}-R^{799} and displays maximal overall dimensions of 70 Å (height) × 85 Å (width) × 65 Å (depth) (Figures 4A and 4B). It consists of four domains: the N-terminal pro-peptide (N^{50}–E^{133}), a metzincin-type CD (Q^{134}–A^{442}), a VEG domain (K^{443}–P^{480} + H^{644}–R^{799}; for acronym explanation, see below), and MAM (a domain previously described in meprin α and β, A5 protein, and receptor-type tyrosine-protein phosphatase μ) (G^{481}–A^{643}), grafted into VEG (see Figure 4C).

Pro-peptide and Catalytic Domain of Pro-*Ba*InhA2-E/A

The polypeptide chain starts on the primed side of the active-site cleft with N^{50}, the first of the 85 residues of the pro-peptide. The chain runs in extended conformation along the cleft in the direction of a true substrate for its first four residues (Figures 4D and 4E). Thereafter, the chain features a small globular moiety framed by two helices (α1 and α2; see Figure 1B). After helix α2, the polypeptide runs across the non-primed side of the cleft and reaches helix α3 at its leftmost end. This helix runs downward along the left molecular surface and reaches the final maturation cleavage point (E^{133}–Q^{134}) in the lower back part of the molecule, on top of the VEG domain (see below and Figure 4D). This segment, including α3, is flexible and mostly absent

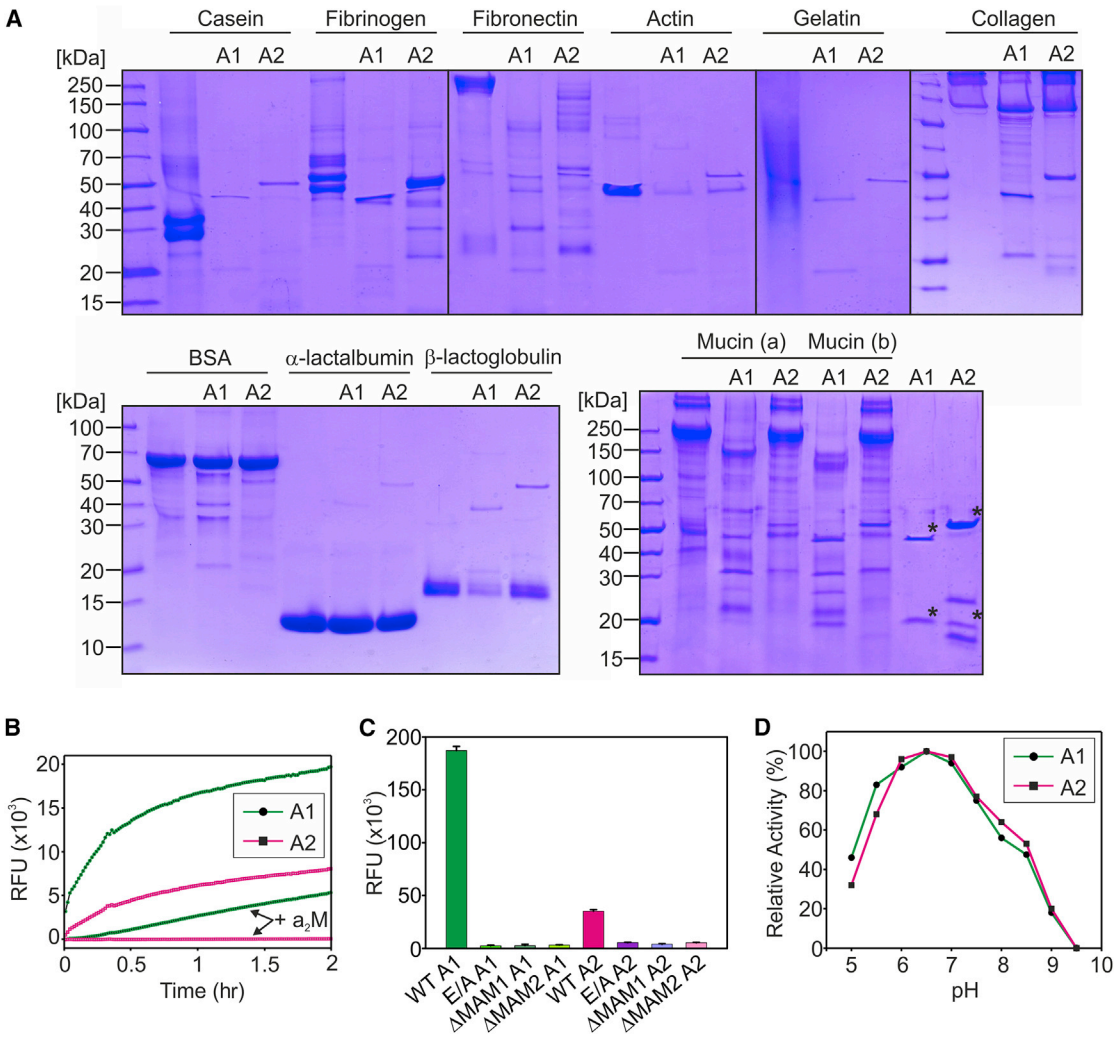


Figure 3. Proteolytic Activity of *Ba*InhA1 and *Ba*InhA2

(A) Protein substrates in the absence and presence of mature WT *Ba*InhA1 or *Ba*InhA2 produced in *B. subtilis*. Reactions were carried out overnight at either 1:10 (collagen) or 1:50 (rest of substrates) enzyme/substrate weight ratios. Mucin was incubated at weight ratios of 1:50 (a) and 1:10 (b). Asterisks indicate bands corresponding to final autolytic products of the peptidases.

(B) Proteolysis of fluorogenic casein over time by *Ba*InhA1 (at 0.01 μg/ml) or *Ba*InhA2 (0.1 μg/ml) produced in *B. subtilis*. Both proteins were further incubated with human α_2-macroglobulin (α_2M) at an enzyme/inhibitor molar ratio of 1:2.

(C) Proteolytic activity on fluorogenic casein of WT, mutant, and MAM-deleted variants of *Ba*InhA1 and *Ba*InhA2 produced in *B. anthracis* (all at 0.01 μg/ml). Values from triplicate reactions were averaged and standard deviations calculated.

(D) pH Dependence of *Ba*InhA1 and *Ba*InhA2 cleavage activity.

(from G^{89}–E^{95} to P^{132}–L^{143}) in three of the four molecules of the two structures solved.

At Q^{134}, the chain enters the CD (see Figures 4A–4C), which with 309 residues is the longest of any metzincin described to date and is subdivisible into an NTS (Q^{134}–G^{386}) and a CTS (L^{387}–A^{442}; see Figure 4C). The first 22 residues of the NTS creep upward along the back molecular surface interacting with the VEG domain. At R^{162}, the polypeptide enters the first strand (β1) of the metzincin-characteristic five-stranded β sheet (top to bottom, β4+β5↑-β1↑-β6↑-β9↑-β7↓). Downmost strand β7 is antiparallel to the others and creates the upper rim of the active-site cleft, which has a slightly electronegative character (Figure 4E). In metzincins, a backing helix connects the strands topologically equivalent to β1 and β4+β5. Here, this helix is divided into two (α4 and α5) by a protrusion of the polypeptide featuring a β ribbon (strands β2 and β3), which contributes to the right back surface. Additional elements decorate the central NTS β sheet. Among these, an adamalysin helix (α7) nestles on top of the β sheet and connects β4+β5 with β6. Helix α7 is preceded by a short upstream helix α6 and loop (A^{252}–G^{265}), which contributes to a calcium site (site 1) together with Lβ6β7 (G^{317}–S^{334}; Figure 5A). These two loops constitute a bulge, which protrudes from the molecular surface above the cleft on its primed side. Downstream of helix α7, a long loop (G^{281}–H^{310}) includes 3_{10}-helix η1 and shapes the left side of the molecule. This loop, hereafter the calcium-binding loop, is stabilized by two

Table 1. Crystallographic Data

Dataset	*Ba*InhA2-E/A (Native)	*Ba*InhA2-E/A (SeMet, Peak)	*Ba*InhA2-E/A (SeMet, Peak)[a]	*Ba*InhA2-E/A (SeMet, Remote)[a]
Space group	C2	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell constants *a*, *b*, *c* (Å), β (°)	177.0, 108.9, 100.4, 119.1	97.6, 102.4, 242.9, 90.0	97.6, 102.4, 242.9, 90.0	97.2, 102.1, 242.3, 90.00
Wavelength (Å)	0.97934	0.97926	0.97926	0.93923
No. of measurements/unique reflections	348,116/51,133	457,866/54,813	459,055/104,066	731,055/103,236
Resolution range (Å) (outermost shell)[b]	47.7–2.60 (2.74–2.60)	48.8–2.90 (3.07–2.90)	48.8–2.90 (3.07–2.90)	48.6–2.90 (3.06–2.90)
Completeness (%)	99.5 (99.2)	99.8 (99.8)	99.8 (99.8)	100.0 (100.0)
R_{merge}[c]	0.105 (0.956)	0.114 (0.672)	0.097 (0.630)	0.159 (1.075)
R_{rim} (= R_{meas})[d]/$CC_{1/2}$[d]	0.113 (1.036)/0.998 (0.746)	0.121 (0.715)/0.998 (0.942)	0.111 (0.716)/ 0.997 (0.876)	0.172 (1.161)/ 0.996 (0.793)
Average intensity[e]	14.8 (2.1)	15.2 (4.2)	13.4 (3.1)	11.2 (2.1)
B Factor (Wilson) ($Å^2$)/average multiplicity	55.6/6.8 (6.7)	54.4/8.4 (8.5)	54.7/4.4 (4.4)	55.9/7.1 (7.0)
Anomalous Corr.[e]/ SigAno[e]/Nano[e]	–/–/–	–/–/–	47/1.30/49,083	30/1.00/48,942
Resolution range used for refinement (Å)	47.7–2.60	42.5–2.90		
No. of reflections used (test set)	51,130 (729)	54,706 (706)		
Crystallographic *R* factor (free *R* factor)[c]	0.171 (0.211)	0.182 (0.207)		
No. of protein atoms/solvent molecules/neutral (covalent) ligands/ionic ligands	11,399/168/3 acetonitrile/2 Zn^{2+}, 9 Ca^{2+}, 2 K^+, 2 Na^+	10,763/104/2 sulfobetaine, 6 glycerol/2 Zn^{2+}, 8 Ca^{2+}, 2 K^+		
Root-mean-square deviation from target values				
Bonds (Å)/angles (°)	0.010/1.15	0.010/1.14		
Average *B* factors ($Å^2$)	61.2	71.8		
All-atom contacts and geometry analysis[f]				
Residues				
In favored regions/outliers/ all residues	1,400 (96.2%)/ 9 (0.6%)/1,456	1,321 (96.3%)/ 2 (0.1%)/1,372		
With poor rotamers/bad bonds/bad angles	30 (2.23%)/0/1	37 (3.26%)/0/1		
With Cβ deviations >0.25 Å/clashscore	0/5.06 (99th percentile)	1/6.99 (98th percentile)		
MolProbity score	1.79 (99th percentile)	2.02 (99th percentile)		

[a]Friedel mates were treated separately during processing for substructure solution and phasing.
[b]Data processing values in parentheses refer to the outermost resolution shell.
[c]For definitions, see Table 1 in García-Castellanos et al. (2003).
[d]For definitions, see Evans, 2006; Karplus and Diederichs, 2012.
[e]Average intensity is $\langle I/\sigma(I)\rangle$ of unique reflections after merging, according to the XDS and XSCALE programs (Kabsch, 2010).
[f]According to MolProbity (Chen et al., 2010).

vicinal calcium cations separated by 10 Å (site 2; see Figure 5B), which is reminiscent of the double calcium site found in the papain family member ulilysin within metzincins (Tallant et al., 2006). Stability provided by calcium was suggested for *B. thuringiensis* InhA1 (Lövgren et al., 1990) and was verified here for *Ba*InhA2 by thermal-shift experiments (Figure 2F). The calcium-binding loop and the adamalysin helix surround loop Lβ7β9, which includes short strand β8 interacting with the beginning of β9, on the upper surface of the molecule. After strand β9, the chain enters active-site helix α8, which runs right to left and frames the active-site cleft on its back. Helix α8 contains most of the extended metal-binding motif, which includes zinc-binding residues H^{379} and H^{383}, and the catalytic glutamate, E^{380}, here mutated to alanine (see Figure 5F). The NTS-CTS transition is G^{386}–L^{387}, and the third zinc-binding residue is D^{389} from the CTS. Most metzincins have a histidine as third zinc ligand, but

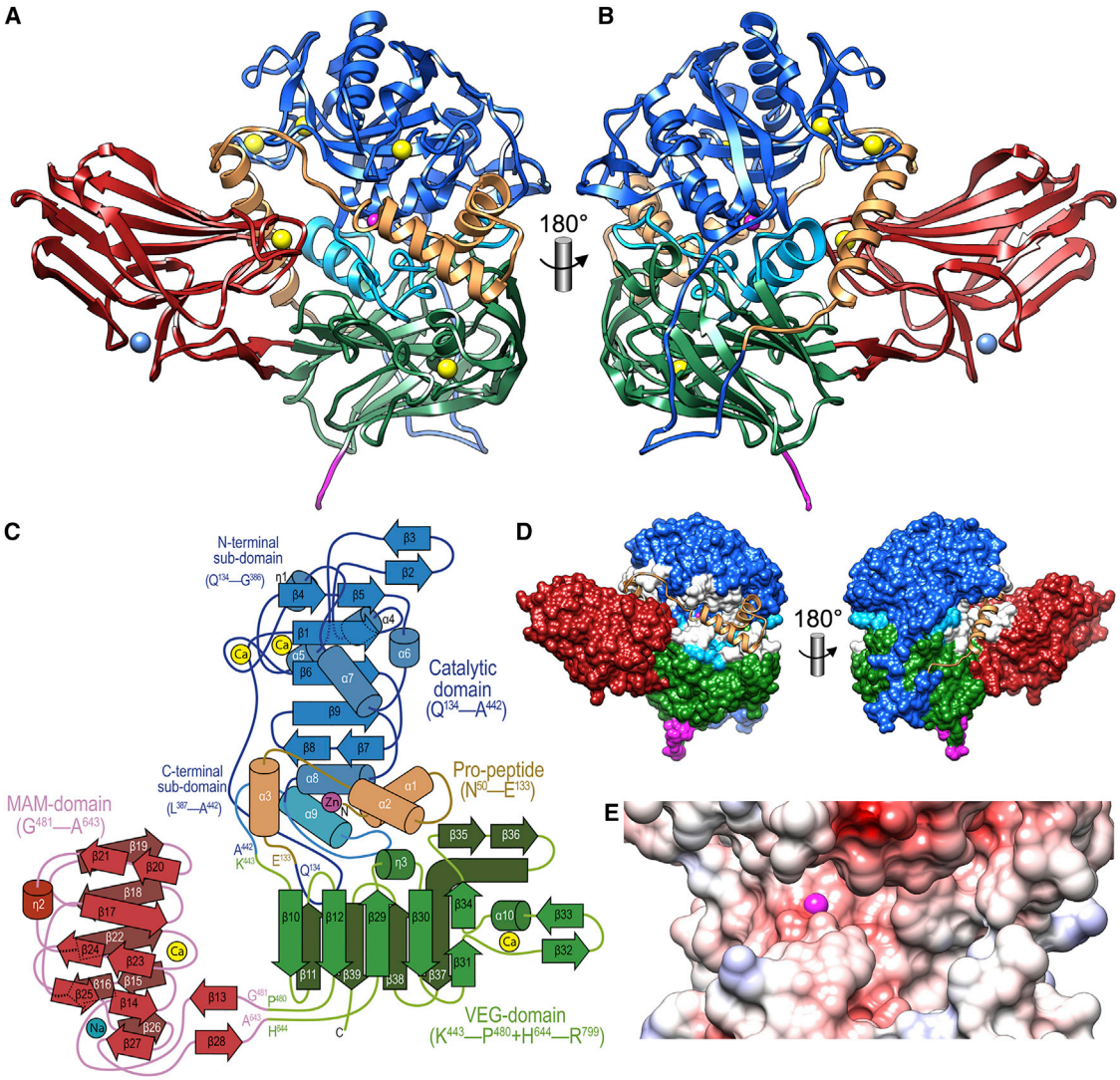


Figure 4. Crystal Structure and Topology of Pro-*Ba*InhA2-E/A

(A and B) Structure in front view (A) and rear view (B) according to the characteristic standard orientation of MPs. The pro-peptide is in orange, the catalytic domain in blue (the NTS in dark blue, the CTS in cyan), the VEG domain in green, and the MAM domain in red. The C-terminal His_6-tag is shown as a magenta ribbon. Calcium, zinc, and sodium ions are shown as yellow, magenta, and blue spheres, respectively.

(C) Topology scheme of pro-*Ba*InhA2-E/A showing the repetitive secondary structure elements and domain boundaries. Colors as in (A).

(D) Pro-*Ba*InhA2-E/A in front and rear views as surface model for the mature enzyme moiety (domains colored as in A) and as a ribbon for the pro-peptide. The regions of the former contributing to the interaction surface are shown in white.

(E) Close-up of the left panel of (D) showing the active-site cleft of the mature protein moiety superimposed with its electrostatic surface, ranging from −10 kcal/(mol·*e*) (red) over 0 kcal/(mol·*e*) (white) to +10 kcal/(mol·*e*) (blue). The MAM domain and the pro-peptide have been omitted for clarity.

See also Figure S2.

aspartates have also been described for snapalysins and igalysins (see Cerdà-Costa and Gomis-Rüth, 2014 and references therein). Some metzincins such as snapalysin and igalysin have a tyrosine two positions downstream of the third ligand, which may play a role in substrate binding and catalysis. It is also present in pro-*Ba*InhA2-E/A as Y^{391}. Overall, the CTS is much smaller than the NTS (56 versus 253 residues) and encompasses a Met-turn typical for metzincins—centered on M^{408}—and a C-terminal helix (α9), which leads to the end of the CD at A^{442}. Similarity searches with *Ba*InhA2 CD identified the MPs ADAM-17 (PDB: 3CKI; Maskos et al., 1998; Figure S2A) and ulilysin (PDB: 2CKI; Tallant et al., 2006) as closest structural relatives, in agreement with their common ascription to the metzincins (Cerdà-Costa and Gomis-Rüth, 2014).

The VEG Domain of Pro-*Ba*InhA2-E/A

After leaving the CD, the polypeptide enters an elaborate β domain, which is reminiscent of the principal viral envelope glycoprotein from dengue virus (PDB: 3UC0; Cockburn et al., 2012; see Figure S2B) and is hereafter dubbed the VEG domain

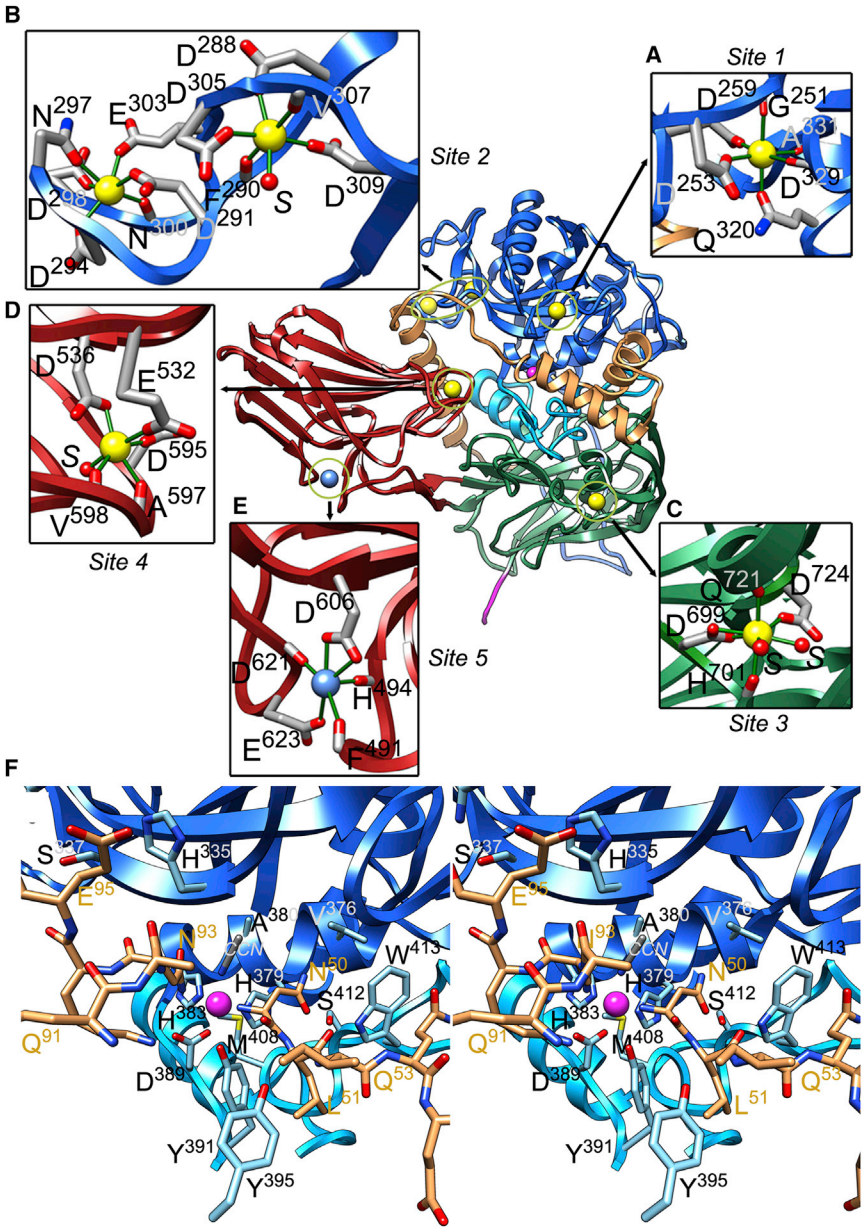


Figure 5. Cation Sites and Active-Site Cleft of Pro-*Ba*InhA2-E/A

If not otherwise stated, calcium sites are octahedral, with oxygens at liganding distances of 2.4 ± 0.1 Å (typical oxygen-binding distances for calcium span 2.36–2.39 Å (Harding, 2006).

(A) Site 1 calcium ion.

(B) Site 2 calcium ions.

(C) Site 3 calcium.

(D) Site 4 calcium. D^{536} also binds through its second carboxylate oxygen, at 2.6 Å.

(E) Sodium cation within the MAM domain octahedrally-1 coordinated by side- and main-chain oxygens of the protein. The binding distances are greater than normal for sodium (2.5–3.3 Å versus 2.38–2.41 Å on average, see Harding, 2006), so it could be a heavier cation with partial occupancy. The site is absent from the two molecules of the SeMet derivative.

(F) Close-up view in cross-eye stereo of the active site of *Ba*InhA2-E/A centered on the catalytic zinc, with the CD as a ribbon colored according to Figures 4A and 4B (residue labels in black/light gray) and two fragments of the pro-peptide as stick models with carbons in orange (labels in orange). Selected side chains are depicted. The zinc is bound by H^{379}Nε2, H^{383}Nε2, D^{389}Oδ1, N^{50}N from the pro-peptide, and a tentative acetonitrile (CCN) nitrogen from the crystallization buffer at distances of 2.1 ± 0.1 Å (typical non-sulfur ligand-binding distances for zinc span 1.99–2.07 Å; Harding, 2006). N^{93}Nδ2 from the pro-peptide is somewhat farther apart (3.1 Å).

See also Figure S2.

(K^{443}–P^{480} + H^{644}–R^{799}). It consists of a central parallel β sandwich (see Figures 4A–4C) made up of an upper five-stranded mixed sheet (β10↓-β12↓-β29↑-β30↓-β31+β34↑) and a lower four-stranded antiparallel sheet (β11↑-β39↓-β38↑-β37↓). This central sandwich is decorated with a β ribbon, perpendicularly grafted after the rightmost strand of the upper sheet and created by β35+β36 with the beginning of strand β37 from the lower sheet. In addition, β35+β36 is interrupted by an insertion that comprises helix α10, a β ribbon (β32β33), and a calcium site (site 3; see Figure 5C). In addition, 3_{10}-helix η3 is grafted onto the upper sheet between the second and third strands from the right, and loop Lη3β30 might contribute to shaping the rightmost edge of the active-site cleft. The VEG domain is attached to the bottom of the CTS (see Figures 4A–4C), mainly through its upper sheet, through an interface spanning 1,610 Å^2. In addition, VEG contacts the N-terminal segment (see above) and the end of α5 plus the downstream loop of NTS.

The MAM Domain of Pro-BaInhA2-E/A

Grafted between the second and third strands of the upper sheet of VEG (Figure 4C), a second β domain is laterally attached. It displays closest structural similarity with MAM domains, such as the one from the human MP, meprin β (PDB: 4QWM; Arolas et al., 2012; see Figure S2C), which in this case is grafted between two domains. The pro-*Ba*InhA2-E/A MAM domain is a β sandwich consisting of two five-stranded antiparallel β sheets rotated away from each other by ~25° (see Figures 4A–4C). The sandwich consists of a back twisted sheet (β19-β18-β22-β15+β16-β26) and a front twisted and curled sheet (β20+β21-β17-β23+β24-β25+β14-β27). It is decorated by 3_{10}-helix η2 after β21, a calcium site (site 4) embraced by Lβ17β18 and Lβ22β23 (see Figure 5D), and a β ribbon (β13β28) included in the segments in extended conformation connecting the MAM and VEG domains. In addition, a tentative sodium site (site 5) is shaped by Lβ13β14, Lβ23β24, and Lβ25β26 (see Figure 5E). While the overall architecture—including the sodium sites—of

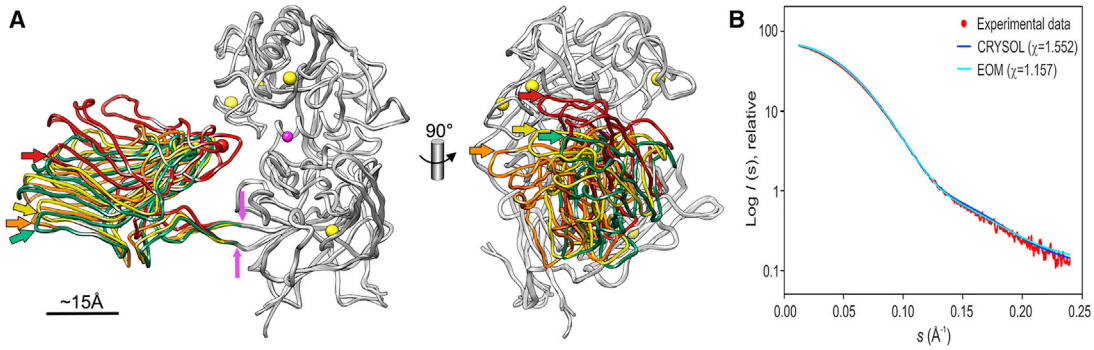


Figure 6. Flexibility in Pro-*Ba*InhA2-E/A

(A) Overlay of the structures of molecules A and B of native *Ba*InhA2 and of the SeMet-derivatized protein in front (left) and lateral (right) views after superposition of the CD and VEG domains (gray ribbons) to highlight the different relative arrangements of the MAM domains (in red, yellow, orange, and green, respectively). Pro-peptides have been omitted for clarity. Equivalent points of the MAM domains are pinpointed by arrows for comparison. Magenta arrows indicate the predicted hinge points, which coincide with the domain limits between VEG and MAM. Magenta, yellow, and red spheres pinpoint the catalytic zinc, calcium, and sodium ions, respectively.

(B) The scattering-intensity curves calculated for the crystal structure of *Ba*InhA2-E/A mutant (present work) using program CRYSOL and expanded for its N-terminal region (20 random conformers out of 1,000 models) using the program EOM are shown in blue and cyan, respectively.

pro-*Ba*InhA2-E/A MAM and meprin β MAM are conserved, their respective insertion into the overall multi-domain moieties is circularly permutated and occurs through different regions (see Figure S2C).

Function of the MAM Domain

MAM domains were initially predicted by bioinformatics, and although their function is unknown, they have been found in cell-surface eukaryotic proteins only and may function as interaction or adhesion domains (Beckmann and Bork, 1993). To explore their role in thuringilysins, we designed several MAM-deleted variants of *Ba*InhA1 and *Ba*InhA2 (see Table S1). Interestingly, all variants were insoluble or not expressed at all in production assays with *E. coli* and *B. subtilis* as hosts, respectively. Production in *B. anthracis* to more closely mimic their natural environment yielded forms that lacked most of the C-terminal VEG domains and were inactive (Figures 3C and S1A–S1C). Removal of these C-terminal parts is likely caused by heterolytic cleavage, since the *B. anthracis* strain used for protein expression lacks MPs, but not proteases, from other catalytic types (Pomerantsev et al., 2011). Overall these results implicate the MAM domain in folding and/or stability of InhA peptidases, which affects susceptibility to (heterolytic) proteolysis.

Flexibility of the VEG-MAM Linkage

Analysis of molecular flexibility of native and SeMet-derivatized pro-*Ba*InhA2-E/A (respective molecules A and B) based on the elastic network model indicated two independent rigid bodies for the pro-peptide + CD + VEG and the MAM moieties separately. P^{480}–G^{481} and A^{643}–H^{644} were identified as hinge points, around which rearrangement through rigid-body rotation would occur. Superposition of the four crystallographic molecules based on the respective CD + VEG moieties revealed that MAM domains are rearranged, but only maximally for ~15 Å in both the horizontal and the vertical planes (see Figure 6A). Maximal displacement in the vertical plane was observed for molecule A of the native structure. This correlated with the segment connecting the globular domain of the pro-peptide with the beginning of the CD being mostly ordered in the final model, in contrast to the other three molecules (see above).

The overall flexibility of the VEG-MAM junction was further analyzed in solution by SAXS using *Ba*InhA1-HE/AA. This was the only full-length protein variant assayed that was stable enough over time, as assessed by SDS-PAGE (see Figure 2A) and dynamic light scattering (DLS), and revealed a homogeneous protein solution (polydispersity <15%). The SAXS experimental curve indicated an essentially globular protein, and best fitting of the experimental data was obtained when the N terminus of the crystallized protein (E^{33}–I^{52}) was modeled as completely flexible (see Figure 6B). The fit, however, did not improve significantly when including the relative arrangement between VEG and MAM as found in the four crystal structures. Accordingly, the hinge between these domains accounts for a certain degree of rotational freedom but not complete flexibility.

Zymogenic and Activity Determinants in Pro-*Ba*InhA2-E/A

Latency is maintained in pro-*Ba*InhA2-E/A through the pro-peptide, which wraps around and interacts with the mature enzyme moiety through a large surface of 3,450 $Å^2$ (Figure 4D). Regions involved include practically the entire pro-peptide and Lα6α7, Lη1β6, Lβ6β7, β7, Lβ7β8, Lα8α9, and α9 of the CD; Lβ14β15 and β22 of the MAM domain; and β36 and the following loop of the VEG domain. The globular moiety of the pro-peptide is attached on the right side of the active-site cleft (see Figures 4A and 4C) and introduces the preceding segment N^{50}–Q^{53} in the primed side of the cleft, thus mimicking a left-behind cleavage product and blocking substrate access (see Figure 5F). The α-amino group of N^{50} contacts the catalytic zinc ion by a zinc-NH_2 coordinative bond. With this ligand and the three zinc ligands from the extended metal-binding motif (see above) plus a tentative acetonitrile molecule from the crystallization solution, the metal adopts a trigonal-bipyramidal coordination. The N^{50} main chain is further bound through a double interaction with $Y^{391}O\eta$, which substantiates a possible role in substrate binding during catalysis as in related MPs (see above). The side chain of

N^{50}, in turn, points into the S_1' pocket, which is generally a structural element of substrate selectivity in MPs. Here, the pocket is wide and framed by segment G^{410}–W^{413} from the segment following the Met-turn, by downstream A^{418}–G^{419}, and by A^{373}–H^{379} from the active-site helix. The bottom of the cleft is occupied by S^{412}, which binds the side-chain Oδ1 atom of N^{50}. Overall, the pocket is shaped to potentially accommodate a variety of side chains, which is compatible with its broad specificity (see above).

Downstream of N^{50}, the main-chain carbonyl of L^{51} binds the side-chain Nε1 atoms of W^{413}, the amide nitrogen of I^{52} binds Y^{395}Oη, and the side chain of Q^{53} is sandwiched between the W^{413} and Y^{187} side chains. A second region of intimate contact between the CD and the pro-peptide is provided by segment G^{89}–P^{98} of the latter. In particular, N^{93} approaches the metal-binding site interacting with R^{336}N from cleft upper-rim strand β7 through its Oδ1 atom, and both zinc-ligand D^{389} atom Oδ2 and the catalytic metal through its Nδ2 atom (see Figure 5F). Downstream pro-peptide residue E^{95} is trapped between the side chains of H^{258}, H^{335}, and R^{267}.

Altogether, this mechanism of latency through the N terminus of the zymogen is novel for MPs, and is distantly reminiscent of the inhibition mechanism of the tissue inhibitors of metalloproteinases (Gomis-Rüth et al., 1997) and serralysin inhibitors (Baumann et al., 1995), which are likewise both inhibited by the N-terminal segment of the inhibitors on the primed side of the respective clefts.

Conclusions

The present study documents a novel mechanism of latency maintenance in MPs through the N terminus, which may be shared by other BCG orthologs and paralogs from the thuringilysin family. Sequence analysis of the slightly more distant *Vc*PrtV (Figure S3) in the light of the structure of pro-*Ba*InhA2-E/A (Figure S3), together with the recently determined NMR structure of its isolated pro-domain (Edwin et al., 2015), point to a very similar general architecture in *Vibrio* thuringilysins. They possess two additional C-terminal domains in tandem called PKD1 and PKD2, which are believed to participate in protein secretion (Rompikuntal et al., 2015) and should not interfere with the preceding domains (see Figure S3). Accordingly, our conclusions on structure, function, and activation of *Ba*InhAs may be valid for these relatives too.

InhA and PrtV thuringilysins evince a grafted MAM domain that has a role in proper expression and possibly also in protein function. So far, such domains have been found only in eukaryotes, which suggests that BCG bacteria could have co-opted one through horizontal gene transfer (HGT) during their intimate coexistence with mammalian hosts and integrated it to furnish thuringilysin peptidases. Integration would have occurred under circular permutation of the point of insertion of the MAM domain into preceding/succeeding domains in the respective multi-domain structures. Similar origins through eukaryotic-to-prokaryotic HGT have been postulated for the bacterial MPs fragilysin and karilysin (Cerdà-Costa et al., 2011; Goulas et al., 2011).

EXPERIMENTAL PROCEDURES

Cloning, Expression, and Purification

The sequences of *Ba*InhA1 and *Ba*InhA2 were amplified from genomic *B. anthracis* DNA and cloned into vectors pHT43 and pSW4 for extracellular overexpression in *B. subtilis* WB800N and *B. anthracis* BH460 cells, respectively (for details on procedures and constructs, see Supplemental Experimental Procedures and Table S1). Proteins were therefore secreted into the extracellular medium (typically, Luria-Bertani), which contained sufficient amounts of zinc and calcium to warrant structural and functional competence of the MPs. E/A and HE/AA mutants of both proteins were prepared using the QuikChange Site-Directed Mutagenesis kit and produced as the WT forms. The SeMet variant of *Ba*InhA2-E/A was obtained in the same way, except that cells were grown in minimal medium supplemented with SeMet. Proteins produced in *B. subtilis* and used for structural assays were purified from the culture supernatant by two consecutive steps of immobilized metal-affinity chromatography using first nickel-nitrilotriacetic acid resin and then TALON resin. Proteins were finally polished by SEC in 20 mM Tris-HCl and 150 mM NaCl (pH 7.5), and concentrated by ultrafiltration using a 10-kDa cutoff filter device.

Autolytic, Proteolytic, and Inhibitory Activity

Autolytic activation of WT and mutant *Ba*InhA1 and *Ba*InhA2 was monitored over time either directly on culture supernatants or after metal-affinity purification by SDS-PAGE and Edman sequencing. Limited proteolysis reactions of *Ba*InhA1-and *Ba*InhA2-E/A mutants were performed using trypsin, chymotrypsin, subtilisin, and thermolysin, and subsequently analyzed by SDS-PAGE. Cleavage activity of mature WT *Ba*InhA1 and *Ba*InhA2 was measured against fluorescein conjugates (BODIPY FL casein, DQ gelatin, and DQ BSA), natural protein substrates (actin, gelatin, collagen type I, mucin, fibronectin, fibrinogen, casein, albumin, α-lactalbumin, and β-lactoglobulin), and ten different fluorogenic peptidic substrates. The optimal pH for activity was determined for both mature WT proteins using BODIPY FL casein as substrate in MES (2-(*N*-morpholino)ethanesulfonic acid) buffer for pH 5.0–7.0; HEPES for pH 6.0–8.0; Tris for pH 7.0–9.0; and CAPS (*N*-cyclohexyl-3-aminopropanesulfonic acid) for pH 8.0–10, respectively. Inhibition was assessed with a battery of inhibitors of different protease classes, as well as the universal pan-protease inhibitor α_2-macroglobulin, and BODIPY FL casein as substrate (see Supplemental Experimental Procedures and Table S2).

Conformational Stability, DLS, and SAXS Measurements

Mature WT *Ba*InhA1 and *Ba*InhA2 were analyzed by the thermal-shift approach in the absence and presence of either 2 mM calcium chloride or 2 mM EDTA to determine their temperatures of mid-transition. The conformation of both proteins, as well as of MAM-deleted variants, was assessed at 25°C by far-UV circular dichroism spectroscopy. Thermal stability of the samples was additionally investigated by heating from 25°C to 95°C over time. *Ba*InhA1-HE/AA was characterized by DLS following standard procedures. SAXS data of this double mutant were collected at 20°C at beamline BM29 (European Synchrotron Radiation Facility [ESRF]) and analyzed following standard procedures.

Crystallization and Structure Determination

Native and SeMet variants of zymogenic *Ba*InhA2-E/A were crystallized at 4°C by the hanging-drop vapor diffusion method with 0.1 M Bis-Tris propane, 0.2 M sodium/potassium tartrate, and 20% (w/v) polyethylene glycol 3350 (pH 7.5) as precipitating-agent solution. Crystals were cryo-protected with 20% (v/v) glycerol, and diffraction datasets were collected at ESRF and ALBA synchrotrons. The crystal structure of the SeMet variant was solved by two-wavelength anomalous diffraction using a dataset collected at the absorption peak wavelength of selenium (determined by a previous XANES scan) and a far-energy remote dataset from the same crystal. The structure of native zymogen was solved by likelihood-scored molecular replacement. Full details on data collection, processing, structure determination, and refinement are given in Supplemental Experimental Procedures and Table 1.

ACCESSION NUMBERS

The atomic coordinates and structure factors of the SeMet and native variants of pro-*Ba*InhA2-E/A have been deposited with the PDB under codes PDB: 4YU5 and 4YU6, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, two tables, and three figures and can be found with this article online at http://dx.doi.org/10.1016/j.str.2015.10.015.

AUTHOR CONTRIBUTIONS

Experiments were conceived and designed by J.L.A., S.H.L., and F.X.G.-R. Reagents and analytical tools were provided by S.H.L. and F.X.G.-R. Experiments and data analysis were performed by J.L.A., T.G., A.P.P., S.H.L., and F.X.G.-R. The paper was written by J.L.A., S.H.L., and F.X.G.-R. with input from T.G. and A.P.P.

ACKNOWLEDGMENTS

We thank Tibisay Guevara, Joan Pous, and Xandra Kreplin (the latter two from the IBMB/IRB Protein Crystallography Platform) for excellent technical assistance during crystallization experiments, and Mar López Pelegrín and Isabel Usón for help with purification experiments and diffraction data collection, respectively. Robin Rycroft is thanked for critical reading of the manuscript. We are indebted to Euripedes de Almeida Ribeiro and Kristina Djinovic-Carugo (both at University of Vienna, Austria), as well as Pau Bernadó (University of Montpellier, France) for excellent guidance on SAXS data analysis. We are also grateful to Hideaki Nagase (University of Oxford, UK), Gunnar Hansson (University of Gothenburg, Sweden), and Lars Sottrup-Jensen (Aarhus University, Denmark) for generous gifts of collagen, mucin, and α_2M, respectively. This study was supported in part by grants from European, Spanish, and Catalan agencies (FP7-PEOPLE-2011-ITN-290246 "RAPID"; FP7-HEALTH-2012-306029-2 "TRIGGER"; BFU2012-32862; BFU2015-64487R; MDM-2014-0435; BIO2013-49320-EXP; and 2014SGR9), and by the intramural program of the National Institute of Allergy and Infectious Diseases, NIH. The Department of Structural Biology of IBMB is a "María de Maeztu" Unit of Excellence of the Spanish Ministry of Economy and Competitiveness. We acknowledge the help provided by ESRF and ALBA synchrotron local contacts. Funding for traveling and synchrotron data collection was provided in part by ESRF.

Received: September 6, 2015
Revised: October 15, 2015
Accepted: October 15, 2015
Published: January 5, 2016

REFERENCES

Arolas, J.L., Broder, C., Jefferson, T., Guevara, T., Sterchi, E.E., Bode, W., Stöcker, W., Becker-Pauly, C., and Gomis-Rüth, F.X. (2012). Structural basis for the sheddase function of human meprin β metalloproteinase at the plasma membrane. Proc. Natl. Acad. Sci. USA *109*, 16131–16136.

Artenstein, A.W., and Opal, S.M. (2012). Novel approaches to the treatment of systemic anthrax. Clin. Infect. Dis. *54*, 1148–1161.

Baillie, L., and Read, T.D. (2001). *Bacillus anthracis*, a bug with attitude! Curr. Opin. Microbiol. *4*, 78–81.

Baumann, U., Bauer, M., Letoffe, S., Delepelaire, P., and Wandersman, C. (1995). Crystal structure of a complex between *Serratia marcescens* metalloprotease and an inhibitor from *Erwinia chrysanthemi*. J. Mol. Biol. *248*, 653–661.

Beckmann, G., and Bork, P. (1993). An adhesive domain detected in functionally diverse receptors. Trends Biochem. Sci. *18*, 40–41.

Beddoe, T., Paton, A.W., Le Nours, J., Rossjohn, J., and Paton, J.C. (2010). Structure, biological functions and applications of the AB_5 toxins. Trends Biochem. Sci. *35*, 411–418.

Cerdà-Costa, N., and Gomis-Rüth, F.X. (2014). Architecture and function of metallopeptidase catalytic domains. Protein Sci. *23*, 123–144.

Cerdà-Costa, N., Guevara, T., Karim, A.Y., Ksiazek, M., Nguyen, K.A., Arolas, J.L., Potempa, J., and Gomis-Rüth, F.X. (2011). The structure of the catalytic domain of *Tannerella forsythia* karilysin reveals it is a bacterial xenologue of animal matrix metalloproteinases. Mol. Microbiol. *79*, 119–132.

Charlton, S., Moir, A.J., Baillie, L., and Moir, A. (1999). Characterization of the exosporium of *Bacillus cereus*. J. Appl. Microbiol. *87*, 241–245.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Chitlaru, T., Gat, O., Gozlan, Y., Ariel, N., and Shafferman, A. (2006). Differential proteomic analysis of the *Bacillus anthracis* secretome: distinct plasmid and chromosome CO_2-dependent cross talk mechanisms modulate extracellular proteolytic activities. J. Bacteriol. *188*, 3551–3571.

Chung, M.C., Popova, T.G., Millis, B.A., Mukherjee, D.V., Zhou, W., Liotta, L.A., Petricoin, E.F., Chandhoke, V., Bailey, C., and Popov, S.G. (2006). Secreted neutral metalloproteases of *Bacillus anthracis* as candidate pathogenic factors. J. Biol. Chem. *281*, 31408–31418.

Chung, M.C., Jorgensen, S.C., Tonry, J.H., Kashanchi, F., Bailey, C., and Popov, S. (2011). Secreted *Bacillus anthracis* proteases target the host fibrinolytic system. FEMS Immunol. Med. Microbiol. *62*, 173–181.

Cockburn, J.J.B., Navarro Sanchez, M.E., Goncalvez, A.P., Zaitseva, E., Stura, E.A., Kikuti, C.M., Duquerroy, S., Dussart, P., Chernomordik, L.V., Lai, C.-J., and Rey, F.A. (2012). Structural insights into the neutralization mechanism of a higher primate antibody against dengue virus. EMBO J. *31*, 767–779.

Edlund, T., Sidén, I., and Boman, H.G. (1976). Evidence for two immune inhibitors from *Bacillus thuringiensis* interfering with the humoral defense system of saturniid pupae. Infect. Immun. *14*, 934–941.

Edwin, A., Persson, C., Mayzel, M., Wai, S.N., Ohman, A., Karlsson, B.G., and Sauer-Eriksson, A.E. (2015). Structure of the N-terminal domain of the metalloprotease PrtV from *Vibrio cholerae*. Protein Sci. *24*, http://dx.doi.org/10.1002/pro.2815.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. *62*, 72–82.

García-Castellanos, R., Marrero, A., Mallorquí-Fernández, G., Potempa, J., Coll, M., and Gomis-Rüth, F.X. (2003). Three-dimensional structure of MecI : molecular basis for transcriptional regulation of staphylococcal methicillin resistance. J. Biol. Chem. *278*, 39897–39905.

Gomis-Rüth, F.X. (2003). Structural aspects of the metzincin clan of metalloendopeptidases. Mol. Biotechnol. *24*, 157–202.

Gomis-Rüth, F.X. (2013). A different look for AB_5 toxins. Structure *21*, 1909–1910.

Gomis-Rüth, F.X., Maskos, K., Betz, M., Bergner, A., Huber, R., Suzuki, K., Yoshida, N., Nagase, H., Brew, K., Bourenkov, G.P., et al. (1997). Mechanism of inhibition of the human matrix metalloproteinase stromelysin-1 by TIMP-1. Nature *389*, 77–81.

Goulas, T., Arolas, J.L., and Gomis-Rüth, F.X. (2011). Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. Proc. Natl. Acad. Sci. USA *108*, 1856–1861.

Guillemet, E., Cadot, C., Tran, S.L., Guinebretiere, M.H., Lereclus, D., and Ramarao, N. (2010). The InhA metalloproteases of *Bacillus cereus* contribute concomitantly to virulence. J. Bacteriol. *192*, 286–294.

Harding, M.M. (2006). Small revisions to predicted distances around metal sites in proteins. Acta Crystallogr. D Biol. Crystallogr. *62*, 678–682.

Johansson, M.E., Sjövall, H., and Hansson, G.C. (2013). The gastrointestinal mucus system in health and disease. Nat. Rev. Gastroenterol. Hepatol. *10*, 352–361.

Kabsch, W. (2010). XDS. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030–1033.

Koch, R. (1876). Die Aetiologie der Milzbrand-Krankheit begründet auf die Entwicklungsgeschichte des *Bacillus anthracis*. Cohn's Beitr. Z. Biol. D Pflanz *2*, 277–310.

Lövgren, A., Zhang, M., Engström, A., Dalhammar, G., and Landén, R. (1990). Molecular characterization of immune inhibitor A, a secreted virulence protease from *Bacillus thuringiensis*. Mol. Microbiol. *4*, 2137–2146.

Maskos, K., Fernández-Catalán, C., Huber, R., Bourenkov, G.P., Bartunik, H., Ellestad, G.A., Reddy, P., Wolfson, M.F., Rauch, C.T., Castner, B.J., et al. (1998). Crystal structure of the catalytic domain of human tumor necrosis factor-α-converting enzyme. Proc. Natl. Acad. Sci. USA *95*, 3408–3412.

McKenzie, A.T., Pomerantsev, A.P., Sastalla, I., Martens, C., Ricklefs, S.M., Virtaneva, K., Anzick, S., Porcella, S.F., and Leppla, S.H. (2014). Transcriptome analysis identifies *Bacillus anthracis* genes that respond to CO_2 through an AtxA-dependent mechanism. BMC Genomics *15*, 229.

Mukherjee, D.V., Tonry, J.H., Kim, K.S., Ramarao, N., Popova, T.G., Bailey, C., Popov, S., and Chung, M.C. (2011). *Bacillus anthracis* protease InhA increases blood-brain barrier permeability and contributes to cerebral hemorrhages. PLoS One *6*, e17921.

Ng, N.M., Littler, D.R., Paton, A.W., Le Nours, J., Rossjohn, J., Paton, J.C., and Beddoe, T. (2013). EcxAB is a founding member of a new family of metalloprotease AB toxins with a hybrid cholera-like B subunit. Structure *21*, 2003–2013.

Pflughoeft, K.J., Swick, M.C., Engler, D.A., Yeo, H.J., and Koehler, T.M. (2014). Modulation of the *Bacillus anthracis* secretome by the immune inhibitor A1 protease. J. Bacteriol. *196*, 424–435.

Pomerantsev, A.P., Pomerantseva, O.M., Moayeri, M., Fattah, R., Tallant, C., and Leppla, S.H. (2011). A *Bacillus anthracis* strain deleted for six proteases serves as an effective host for production of recombinant proteins. Protein Expr. Purif. *80*, 80–90.

Rasko, D.A., Altherr, M.R., Han, C.S., and Ravel, J. (2005). Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol. Rev. *29*, 303–329.

Rompikuntal, P.K., Vdovikova, S., Duperthuy, M., Johnson, T.L., Ahlund, M., Lundmark, R., Oscarsson, J., Sandkvist, M., Uhlin, B.E., and Wai, S. (2015). Outer membrane vesicle-mediated export of processed PrtV protease from Vibrio cholerae. PLoS One *10*, e0134098.

Schutte, A., Ermund, A., Becker-Pauly, C., Johansson, M.E., Rodriguez-Piñeiro, A.M., Backhed, F., Muller, S., Lottaz, D., Bond, J.S., and Hansson, G.C. (2014). Microbial-induced meprin β cleavage in MUC2 mucin and a functional CFTR channel are required to release anchored small intestinal mucus. Proc. Natl. Acad. Sci. USA *111*, 12396–12401.

Sternbach, G. (2003). The history of anthrax. J. Emerg. Med. *24*, 463–467.

Tallant, C., García-Castellanos, R., Seco, J., Baumann, U., and Gomis-Rüth, F.X. (2006). Molecular analysis of ulilysin, the structural prototype of a new family of metzincin metalloproteases. J. Biol. Chem. *281*, 17920–17928.

Tallant, C., García-Castellanos, R., Baumann, U., and Gomis-Rüth, F.X. (2010). On the relevance of the Met-turn methionine in metzincins. J. Biol. Chem. *285*, 13951–13957.

Terwilliger, A., Swick, M.C., Pflughoeft, K.J., Pomerantsev, A., Lyons, C.R., Koehler, T.M., and Maresso, A. (2015). *Bacillus anthracis* overcomes an amino acid auxotrophy by cleaving host serum proteins. J. Bacteriol. *197*, 2400–2411.

Tonry, J.H., McNichol, B.A., Ramarao, N., Chertow, D.S., Kim, K.S., Stibitz, S., Schneewind, O., Kashanchi, F., Bailey, C.L., Popov, S., and Chung, M.C. (2012). *Bacillus anthracis* protease InhA regulates BslA-mediated adhesion in human endothelial cells. Cell. Microbiol. *14*, 1219–1230.

Trillo-Muyo, S., Martínez-Rodríguez, S., Arolas, J.L., and Gomis-Rüth, F.X. (2013). Mechanism of action of a Janus-faced single-domain protein inhibitor simultaneously targeting two peptidase classes. Chem. Sci. *4*, 791–797.

Vaitkevicius, K., Rompikuntal, P.K., Lindmark, B., Vaitkevicius, R., Song, T., and Wai, S.N. (2008). The metalloprotease PrtV from *Vibrio cholerae*. FEBS J. *275*, 3167–3177.

