Proenzyme Structure and Activation of Astacin Metallopeptidase

Tibisay Guevara, Irene Yiallouros, Reinhold Kappelhoff, Steffen Bisسود, Walter Stöcker, and F. Xavier Gomez-Rüth

From the 4Proteolysis Laboratory, Department of Structural Biology, Molecular Biology Institute of Barcelona, Consejo Superior de Investigaciones Científicas, Barcelona Science Park, Helix Building, c/Baldiri Reixac 15-21, E-08028, Spain and the 4Institute of Zoology, Cell and Matrix Biology, Johannes Gutenberg-University, Johannes-von-Müller-Weg 6, D-55128 Mainz, Germany

Proteolysis is regulated by inactive (latent)zymogens, with a prosegment preventing access of substrates to the active-site cleft of the enzyme. How latency is maintained often depends on the catalytic mechanism of the protease. For example, in several families of the metzincin metallopeptidases, a “cysteine switch” mechanism involves a conserved prosegment motif with a cysteine residue that coordinates the catalytic zinc ion. Another family of metzincins, the astacins, do not possess a cysteine switch, so latency is maintained by other means. We have solved the high resolution crystal structure of proastacin from the European crayfish, Astacus astacus. Its prosegment is the shortest structurally reported for a metallopeptidase, and it has a unique structure. It runs through the active-site cleft in reverse orientation to a genuine substrate. Moreover, a conserved aspartate, projected by a wide loop of the prosegment, coordinates the zinc ion instead of the catalytic solvent molecule found in the mature enzyme. Activation occurs through two-step limited proteolysis and entails major rearrangement of a flexible activation domain, which becomes rigid and creates the base of the substrate-binding cleft. Maturation also requires the newly formed N terminus to be precisely trimmed so that it can participate in a buried solvent-mediated hydrogen-bonding network, which includes an invariant active-site residue. We describe a novel mechanism for latency and activation, which shares some common features both with other metallopeptidases and with serine peptidases.

The proteolytic activity of most metallopeptidases (MPs) is regulated, and it is only exerted where and when required (1). Such control may occur through modulation of gene expression, compartmentalization, allosterry, or inhibition by protein inhibitors. Another regulatory mechanism is zymogenic latency, which is provided by mostly N-terminal prosegments. These block access of substrates to the active-site cleft, and they are removed by limited proteolysis during maturation (2, 3). Such prosegments often fold independently and guide on their part the folding process of the cognate protease domains. They may also act as intramolecular chaperones or inhibitors of the mature enzymes in trans and in intracellular sorting of the zymogen (2). Therefore, the study of the molecular mechanisms by which MPs maintain latency is indispensable to an understanding of their basic mode of action. It also paves the way for the design of inhibitors that mimic the latent state so as to modulate MP activity as part of therapeutic approaches. Detailed three-dimensional structural information can contribute much to this understanding (4). However, among the MPs, only funnelins, lysostaphins, thermolysins, and matrix metalloproteinases (MMPs) have been structurally analyzed for their zymogens. Results reveal that the mature enzyme moieties are already in a competent conformation. Notwithstanding, each group displays a distinct mechanism of latency maintenance (5–11).

The metzincins are a clan of MPs characterized by a consensus sequence responsible for binding of the catalytic zinc ion (CSBZ), HEXXHXX(G/N)XX(H/D), and a conserved methionine-containing turn (Met-turn) (12–16). This clan can be subdivided into several families, including the MMPs. In the latter, the CSBZ encompasses three histidine zinc ligands, the general base/acid glutamate for catalysis, and an essential glycine (15). A total of 23 MMP paralogs are present in humans (15). Most of them comprise a prosegment upstream of the catalytic domain, and, to date, structures of pro-MMP-1, -2, -3, and -9 have been reported (17–21). They show 60–90-residue globular prosegments, which comprise a three-helix bundle that creates a scaffold to place a segment in extended conformation to block the active-site cleft. This segment includes a conserved “cysteine switch” or “Velcro” sequence motif (22–24), PRCGXPD, which runs through the active-site cleft in the opposite direction to a...
substrate bound along the cleft. The cysteine Sγ atom binds the catalytic zinc ion, thus replacing the catalytic solvent molecule present in mature MMPs. The competent conformation of the cysteine switch peptide is maintained by a double salt bridge between the arginine and the aspartate of the motif. It was hypothesized that this mechanism is shared, with some variation, by other metzincin families for which conserved cysteine switch motifs were found or suggested, such as the ADAMs/adamalysins (conserved motif PKMCGV (25, 26)), leishmanolysins (motif HRCIHD (14)), and pappalysins (motif CG (27)).

In addition, the metzincins also include the astacins, which owe their name to the prototypical digestive enzyme, astacin, from the European freshwater crayfish, *Astacus astacus* L. (13, 28–30). Astacins are found throughout the animal kingdom and in bacteria. Several of these enzymes are crucial for embryonic development, tissue differentiation, and extracellular matrix assembly, and they are thus therapeutic targets. They consist of ~200-residue catalytic moieties preceded by prosegments of varying length (Fig. 1). Several family members have additional downstream domains engaged in substrate recognition, membrane anchors, or signaling. The prototypical crayfish astacin is synthesized as a preproenzyme in the midgut gland (31). The 15-residue signal peptide is removed during secretion, and the proenzyme, which comprises a 34-residue prosegment and a 202-residue catalytic domain, is only transiently found within the ducts between the hepatopancreas and the stomach. Once in the stomach, the zymogen is activated, and the mature enzyme participates in collagenolysis and gelatinolysis during digestion (32). Optimal astacin substrates comprise at least five amino acids with small aliphatic residues in P-1, proline in P-2, bulky hydrophobic residues in P-3, and basic residues in P-1 and P-2 (peptide substrate and enzyme subsite nomenclature according to Refs. 33 and 34). Cleavage is best performed at pH 6–8. A comprehensive review on the enzyme was written by Stöcker and Yiallourou (31).

The prosegment of most astacins contains a short sequence with a conserved aspartic acid residue (Fig. 1). However, there is no chemical evidence for a function of this aspartate in latency maintenance; site-directed mutations of this residue to alanine or asparagine and subsequent expression of the mutant proteins in bacteria resulted in non-foldable, unstable proastacin variants (35). Therefore, we solved the high resolution structure of proastacin in an attempt to shed light on this potentially novel mechanism of latency among MPs.

**EXPERIMENTAL PROCEDURES**

**Protein Preparation and Crystallization**—Recombinant proastacin E93MA/I91ML (UniProt Q9U918; numbering is based on the mature enzyme (M); see below) was expressed in *Escherichia coli* BL21(DE3) cells as inclusion bodies, purified by Ni2+–nitrilotriacetic acid affinity chromatography, and folded by dilution and removal of reducing agents and guanidinium chloride as described (32). The resulting protein in 50 mM AMPSO buffer, pH 9.0, was concentrated in Amicon Ultra-15 centrifugal filter units of 10 kDa cut-off (Millipore, Bad Schwalbach, Germany) to a volume of about 1 ml. The protein was applied to a HEMA-Bio gel filtration column (MZ Analysetechnik, Mainz, Germany) attached to a Kontron high pressure liquid chromatography unit to separate monomeric proastacin from an oligomeric fraction with a flow rate of 0.5 ml/min. The composition of the fractions was analyzed by native PAGE and SDS-PAGE. Crystallization assays were performed with the monomeric protein fraction by the sitting drop vapor diffusion method. Reservoir solutions were prepared by a Tecan robot, and 200-nl crystallization drops were dispensed on 96 2-well MRC plates (Innovadyne) by a Cartesian (Genomic Solutions) nanodrop robot at the High-Throughput Crystallography Platform of the Barcelona Science Park. Best crystals appeared in a Bruker steady temperature crystal farm at 4 °C with protein solution (10 mg/ml in 50 mM AMPSO, pH 9.0) and 20% polyethylene glycol 8000, 0.1 M (NH4)2SO4, 0.01 M MgCl2, 0.05 M MES, pH 5.6, as reservoir solution. These conditions were efficiently scaled up to the microliter range with 24-well Cryochem crystallization dishes (Hampton Research). Crystals were cryoprotected with 16% polyethylene glycol 8000, 20% glycerol, 0.1 M (NH4)2SO4, 0.01 M MgCl2, 0.05 M MES, pH 5.6. A complete diffraction data set to 1.45 Å resolution was collected at 100 K (provided by an Oxford Cryosystems 700 series cryostream) from a single liquid N2 flash-cryocooled crystal on a MarCCD detector at beam line ID23-2 of the European Synchrotron Radiation Facility (Grenoble, France) within the Block Allocation Group “BAG Barcelona.” The crystal was primitive orthorhombic, with one molecule per asymmetric unit. Diffraction data were integrated, scaled, merged, and reduced with programs XDS (36) and SCALA (37) within the CCP4 suite of programs (see Table 1).

**Structure Solution and Refinement**—The structure of proastacin was solved by Patterson search methods with the program PHASER (38) by using the coordinates of mature astacin (Protein Data Bank code 1AST) (39, 40) as searching model. A single solution was found at 358.0, 23.2, and 149.4 (α, β, and γ in Eulerian angles) and 0.353, 0.094, and −0.624 (x, y, and z as fractional unit cell coordinates) after rigid body refinement. This solution gave a Z score of 17.8 for the rotation function and 25.5 for the translation function, as well as a final log likelihood gain of 1133. A subsequent density improvement step with ARP/wARP (41) with all data rendered an electron density map that enabled straightforward chain tracing. Subsequently, manual model building on a Silicon Graphics work station with TURBO-Frodo (42) alternated with crystallographic refinement with program REFMAC5 until completion of the model (see Table 1). This model contained all of the protein residues of the prosegment (Ser1P to Gly34P; prosegment residues bear the suffix “P”) and the mature protease moiety (Ala1M–Arg201M; suffix “M”) (see Fig. 1 for the numbering convention) and the catalytic zinc ion. Two segments, Ala26P–Gln29P and Asp129M–Pro135M.

**Sequence Alignment**—Amino acid sequences were aligned by using ClustalX (43) with default parameters and then manually readjusted with GENEDOC (44) in accordance with structural constraints emanating from the present data and from observations by Stöcker et al. (28). Before alignment, each sequence was processed in *silico* with SIGNAL P (available on the World Wide Web) to remove the signal peptides (45).
**TABLE 1**

Crystallographic data

| Parameters | Values |
|------------|--------|
| Space group/cell constants a, b, and c (Å) | P2_1 2_1 2_1/55.91, 63.37, 69.19 |
| Wavelength (Å) | 0.8726 |
| No. of measurements/unique reflections | 313,933/44,227 |
| Resolution range (Å) (outermost shell) | 41.9–1.45 (1.54–1.45) |
| Completeness (%) | 98.8 (98.9) |
| R_{free,a} = (\sqrt{R_{free,b}}/R_{free,c})° | 0.044 (0.474) |
| Average intensity over S.D. (((f_i)/(r_i))) | 27.3 (3.8) |
| B-Collector (Wilson) (Å^2)/Average multiplicity | 14.6/7.1 (6.2) |
| Resolution range used for refinement (Å) | 8–1.45 |
| No. of reflections in working set/in test set | 43,398 (768) |
| Crystallographic R_{calc} (free R_{calc})a | 0.152 (0.180) |
| No. of protein atoms/solvent molecules/ligands/ions | 1,928/298/3 glycerols/1 Zn^{2+}, 1 SO_{4}^{2–} |
| Main chain conformational angle analysis* | |
| Residues in favored regions/outliers/all residues | 226/1/240 |

a For definitions, see Table 1 in Ref. 58.

b Including atoms and residues in alternative conformation.

According to MOLPROBITY (59).

---

**RESULTS AND DISCUSSION**

**Overall Structure of Proastacin**—In order to prevent autolysis, proastacin was recombinantly overexpressed as a mutant, in which the general base/acid glutamate, Glu^{13960}M, had been replaced by alanine to create a catalytically compromised variant (see Ref. 32). The properly folded monomeric fraction of this protein crystallized and rendered diffraction data to 1.45 Å resolution, which yielded a detailed and accurate picture of the zymogen structure of proastacin (Fig. 2A). It is an ellipsoid of overall dimensions 55 × 45 × 28 Å divided into three major regions: an N-terminal prosegment (residues Ser^{1P}–Gly^{34P}), an upper subdomain (Ala^{13960M}–Gly^{99M}), and a lower subdomain (Phe^{13960M}–Arg^{214M}) according to the standard orientation for MPs (Figs. 1 and 2B). The two subdomains, of similar size, constitute the mature enzyme moiety. They resemble the two valves of a clamshell that delimitate a deep narrow horizontal active-site cleft, which explains the preference of astacin for elongated substrates.

**The Mature Enzyme Moiety**—The upper subdomain contains a twisted five-stranded β-sheet (strands β3–β7; Fig. 2B), with all strands except the fourth (β6) parallel to each other and to an atom that is bound in the cleft. As determined for the mature enzyme structure, protein and peptide substrates are bound left-to-right, mainly on the non-primed half of the cleft (i.e. N-terminally of the scissile bond in peptide substrates). Binding involves antiparallel inter-main-chain interactions with β6, which constitutes the upper rim strand of the cleft and forms the lower edge of the upper subdomain (39, 40). The concave face of the β-sheet accommodates two helices, the backing helix (α2) and the active-site helix (α3) (Fig. 2B). The latter comprises the first two zinc-binding histidine residues (His^{92M} and His^{96M}, Fig. 2C) and the general base/acid glutamate (here alanine) of the CSBZ. The end of the helix at Gly^{99M} is also the boundary with the lower subdomain, which contains the third zinc-binding histidine residue, His^{102M}. This subdomain comprises a flexible segment, Asp^{129M}–Gly^{138M} (Fig. 2D; see below) and a few repetitive secondary structure elements, mainly a short two-stranded β-sheet (strands β8 and β9; see Fig. 2B) and a C-terminal helix (α4) at the end of the polypeptide chain. The latter helix has a kink at Tyr^{194M}, which bends it about 50°. The C terminus of the protein resides at the molecular surface, and two disulfide bonds (Cys^{42M}–Cys^{198M} and Cys^{54M}–Cys^{84M}) contribute to the overall stability and rigidity of the molecule (Figs. 1 and 2). A structural hallmark of metzincins is the Met-turn (12–15), which spans Ser^{145M}–His^{148M} in astacin and contains a methionine at position 3 that is fully conserved (even for its side-chain conformation) among all structurally characterized metzincin prototypes (12, 14–16). This conservation may be attributable to the maintenance of the central active-site core structure of these enzymes.

A Novel Prosegment Structure—With 34 residues, proastacin has the shortest metallopeptidase prosegment structurally analyzed to date, and it has a unique fold (5–11). In contrast to other (metallo)peptidases, it does not have an intramolecular chaperone function in astacin, as revealed by refolding experiments with the heterologously expressed recombinant mature protein (32, 48, 49). The prosegment is elongated and stabilized through several intramolecular interactions. It runs across the front surface of the mature enzyme moiety blocking the cleft, which is consistent with previous functional studies (32) (Fig. 2, B and E). The interaction with the catalytic domain buries an interface of 1,580 Å^2 and is based on 75 close contacts (<4 Å), among them 20 hydrogen bonds, two organometallic interactions, and hydrophobic contacts between 27 pairs of residues (see Table 2). The N terminus of the prosegment at Ser^{3P} is anchored to the lower left of the lower subdomain through a solvent-mediated hydrogen bond with Ile^{135M} O, within the segment connecting the Met-turn and the C-terminal helix. The polypeptide extends up along the molecular surface to the active-site cleft. From Glu^{6P} on, the chain runs through the active-site cleft in reverse orientation to a substrate approximately to Arg^{25P} and adopts a helical conformation (helix α1) until Tyr^{11P} (Figs. 1 and 2B). This segment nestles in the primed side of the cleft (i.e. C-terminally of the scissile bond in peptide substrates), approximately until subsite S_1′, which is partially occupied by the side chain of Leu^{11P}. This helical structure may contribute to the prevention of autolysis because substrates usually bind in extended conformation to active-site clefts (2). Possibly as a result of the aforementioned helix and the flexible segment, Asp^{129M}–Gly^{138M}, within the lower subdomain, the two subdomains are more separated in proastacin than in the unbound mature enzyme (Protein Data Bank code 1AST) (40). In contrast, in the complex of mature astacin with a reaction
intermediate analogue, the cleft was narrower than in the unbound enzyme (Protein Data Bank code 1QJI) (50). Among the residues of helix $\alpha_9$ and $\alpha_11$, three alanines (Ala7P, Ala8P, and Ala10P) and Tyr12P interact with protein residues of the mature enzyme moiety (Fig. 2E and Table 2). At Asn14P, which contacts preceding and downstream carbonyl oxygens of the prosegment through its $N_2$ atom, the chain protrudes outward from the protein moiety and enters a wide loop that ends at Asp21P. This loop comprises two subsequent 1,4-turns (Asn14P–Met17P and Met17P–Gly20P), which together with hydrophobic interactions of Met17P with both Lys23P and Trp65M give rise to a scaffold. This is optimized for positioning Asp21P to approach the catalytic zinc ion (Fig. 2E). The loop structure is further stabilized by a hydrophobic interaction between the side chains of Phe18P and Leu24P and by three key electrostatic interactions of the side chain of Glu19P (see Table 2) with the third zinc-binding histidine, His102M, with Thr105M, and with the penultimate residue of the prosegment,Arg125P (Glu19P $O_2$–Arg125P $N_1$, 3.01 Å; see Fig. 2E). This last salt bridge is vaguely reminiscent of a similar arrangement found in MMP prosegments (see above).

Asp21P coordinates the catalytic zinc ion from the top in a bidentate manner. Together with the mature enzyme ligands, this results in a distorted octahedral coordination sphere, which is unusual for zinc (51) (Fig. 2C). The mature protein ligands are the $N_2$ atoms of His92M, His96M, and His102M as well as Tyr149M $O_2$. The $O_2$ atom of Asp21P replaces the zinc-bound solvent molecule in the mature enzyme ($0.5$ Å away).

The side chain of Tyr149M, which undergoes a “tyrosine switch” motion upon substrate binding (31), is closer to the conformation in the unbound mature enzyme than in the bound enzyme, where it stabilizes one of the two $gem$-dioleate oxygens of the binding histidine, His102M, with Thr105M, and with the penultimate residue of the prosegment, Arg125P (Glu19P $O_2$–Arg125P $N_1$, 3.01 Å; see Fig. 2E). This last salt bridge is vaguely reminiscent of a similar arrangement found in MMP prosegments (see above).

Asp21P coordinates the catalytic zinc ion from the top in a bidentate manner. Together with the mature enzyme ligands, this results in a distorted octahedral coordination sphere, which is unusual for zinc (51) (Fig. 2C). The mature protein ligands are the $N_2$ atoms of His92M, His96M, and His102M as well as Tyr149M $O_2$. The $O_2$ atom of Asp21P replaces the zinc-bound solvent molecule in the mature enzyme ($0.5$ Å away).

The side chain of Tyr149M, which undergoes a “tyrosine switch” motion upon substrate binding (31), is closer to the conformation in the unbound mature enzyme than in the bound enzyme, where it stabilizes one of the two $gem$-dioleate oxygens of the binding histidine, His102M, with Thr105M, and with the penultimate residue of the prosegment, Arg125P (Glu19P $O_2$–Arg125P $N_1$, 3.01 Å; see Fig. 2E). This last salt bridge is vaguely reminiscent of a similar arrangement found in MMP prosegments (see above).

Asp21P coordinates the catalytic zinc ion from the top in a bidentate manner. Together with the mature enzyme ligands, this results in a distorted octahedral coordination sphere, which is unusual for zinc (51) (Fig. 2C). The mature protein ligands are the $N_2$ atoms of His92M, His96M, and His102M as well as Tyr149M $O_2$. The $O_2$ atom of Asp21P replaces the zinc-bound solvent molecule in the mature enzyme ($0.5$ Å away).

The side chain of Tyr149M, which undergoes a “tyrosine switch” motion upon substrate binding (31), is closer to the conformation in the unbound mature enzyme than in the bound enzyme, where it stabilizes one of the two $gem$-dioleate oxygens of the binding histidine, His102M, with Thr105M, and with the penultimate residue of the prosegment, Arg125P (Glu19P $O_2$–Arg125P $N_1$, 3.01 Å; see Fig. 2E). This last salt bridge is vaguely reminiscent of a similar arrangement found in MMP prosegments (see above).
tetrahedral reaction intermediate (50). However, the distance of the phenolic oxygen to the cation in proastacin is slightly greater than in the unbound mature structure, 3.01 Å versus 2.53 Å, possibly due to the nearby side chain of Asp<sup>21P</sup> (Fig. 2C).

The four residues after Asp<sub>21P</sub> (Ile<sub>22P</sub>–Arg<sub>25P</sub>) run in extended conformation and in opposite orientation to a substrate along the cleft between subsites S<sub>2</sub> and S<sub>5</sub> (Fig. 2). They bind through two inter-main-chain bonds to strand β6 above the cleft (Asp<sup>21P</sup> O–Ser<sup>66M</sup> N and Ser<sup>66M</sup> O–Lys<sup>23P</sup> N; see Table 2). This reverse orientation of the prosegment in the cleft may contribute to the prevention of self-cleavage and was previously described in cysteine protease and MMP zymogens (2). The polypeptide chain reaches the molecular surface after Arg<sub>25P</sub>, whose side chain is anchored via three interactions with the mature enzyme backbone (Table 2). Thereafter, a flexible stretch from Ala<sup>26P</sup> to Gln<sup>29P</sup> leads to a short β-hairpin structure made up by strands β1 (Ala<sup>31P</sup>–Val<sup>33P</sup>) and β2 (Ala<sup>2M</sup>–Leu<sup>4M</sup>; see Fig. 2B). The tip of the hairpin contains the main maturation cleavage point (Gly<sup>34P</sup>–Ala<sup>1M</sup>), which is completely buried within the molecular structure.

A Novel Two-step Mechanism of Activation—Based on biochemical studies, proastacin activation in the crayfish is understood to be a two-step process entailing removal and degradation of the prosegment (32). In addition, maturation involves conformational changes in the first six residues of mature astacin, Ala<sup>1M</sup>–Asp<sup>6M</sup>, and major rearrangement of segment Asp<sup>129M</sup>–Gly<sup>138M</sup> (Fig. 3, A and B), which is flexible in the zymogen and is hereafter referred to as the “activation domain.” This is analogous to serine proteinases, for which chain segments grouped under this term show statistical or thermal disorder in the zymogens, partially obstruct the substrate-binding cleft, and lead to incompetent
oxynion holes. In these enzymes, the activation domain becomes rigid and functional upon activation (2, 52).

In the first activation step, crayfish trypsin would perform initial cleavages and render a series of premature astacin species (32). This step would be facilitated by the flexibility of Ala26P–Ala1M within the final stretch of the prossegment. The last tryptic cut was shown to occur at Arg32P–Val33P, and this would probably disrupt a salt bridge, Arg32P-Glu19P, which is essential for structural integrity of the prossegment (see above). The arginine residue is not accessible on the molecular surface of the zymogen, but either the previous tryptic cleavages or the flexible activation domain beneath could provide the space and motion for the minor conformational changes required to enable access by trypsin (Fig. 2E). Subsequent removal of most of the prossegment would offer sufficient space for the activation domain to adopt its competent conformation by undergoing a hitherto unseen flap motion of up to 10 Å (measured at Tyr133M-Ca) (Fig. 3B). In addition, the activation domain would become rigid and well defined, as denoted by the electron density and average B-factors in the mature astacin structure (data not shown; see Refs. 39 and 40), in this way contributing to a functional substrate-binding cleft and active site (Fig. 3C). The initial multiple tryptic cleavages are reminiscent of MMP activation, during which trimming clips by other proteases in a so-called “bait region” are observed (53). This contrasts with funnellins and trypsin-like serine proteases, in which the first cut during activation generates the mature N terminus (2, 54).

In the second activation step, the premature astacin variants, which are catalytically active (32), would produce subsequent cleavages, eventually giving rise to the competent N terminus at Ala1M. The final autolytic cleavage at Gly34P-Asp6M, which matches the substrate specificity of astacin (see above), was unambiguously shown to occur in cis (32). Although Ala1M-N is about 12 Å away from the catalytic zinc in the proenzyme structure, the structural rearrangement of the activation domain triggered by the previous tryptic cleavages could easily lead the new provisional N-terminal segment, Arg32P-Asp6M, which is too long to adopt the definitive structure (see below), to protrude from the molecular surface and become accessible for intramolecular cleavage. After the second set of cleavages, segment Ala1M-Asp6M would again insert into the molecular moiety. In MMPs, similar trimming yields the competent N terminus, needed to form a salt bridge with a conserved aspartate of the final C-terminal helix (53, 55).

The second cleavage step is indispensable because the new N terminus is buried within the mature enzyme body until Ile3N in a cavity framed by segments Asp131M-Gln142M, Phe100M-Met107M, and Thr185M-Gln189M (Protein Data Bank code 1AST) (39, 40). Residues from these segments, together with seven solvent molecules and the first N-terminal residues, in an intricate, completely buried hydrogen-bonding network that is key for structural integrity of the enzyme (see Fig. 3 of Ref. 39) and that is incompatible with N-terminally elongated polypeptide chains. In particular, the α-amino group of Ala1M is coordinated by three solvent molecules, which are further bound to Asp131M-O62 and His102M-O2; to Tyr101M-O; and to Gln189M-O1, Arg106M-N1, and Glu103M-O1, respectively. The last interaction is of particular importance because Glu103M is the immediately downstream neighbor of the third zinc-binding histidine. It is unique for and invariant within the astacin family (28). Comparison of zymogen and mature enzyme further reveals that upon cleavage at Gly34P-Asp6M, the main chain must undergo a 180° rotation around the ψ main-chain angle of the new N-terminal residue to establish the aforementioned solvent-mediated interactions (Fig. 3B). This is again reminiscent of serine proteases, in which a salt bridge...
between the newly formed N terminus and the aspartate residue upstream of the catalytic serine renders a functional enzyme (52). In contrast to the latter, however, maintenance of the hydrogen-bonding network of the new N terminus seems to be mainly required for structure and stability in astacins. Mutants of proastacin, in which Glu103M had been replaced with glutamine and alanine, evinced equivalent catalytic efficiency but much lower thermal stability (32). This structural rather than functional importance is supported by the position and conformation of Glu103M being actually superimposable in the zymogen and the mature structures. A similar scenario is conceivable for N-terminally elongated pre forms of astacin, which would bear intrinsic activity to permit the last steps in maturation but would be much less stable than the mature end product. This is corroborated by the finding that N-terminally extended variants of the astacin family member meprin were enzymatically active but thermally labile (32, 56).

CONCLUSIONS
Proastacin maturation entails step-wise degradation of the prosegment to ensure irreversibility and to prevent competitive inhibition of the mature enzyme. Once the enzyme is activated, it is kept in check in vivo by protein inhibitors of the requisite specificity, such as $\alpha$-2-macroglobulin (57). Previously structurally documented MP activation processes, as seen in MMPs, revealed preformed protease moieties, which merely had to be uncovered by removal of the prosegment. By contrast, activation is accompanied by major rearrangement in astacin and probably in the other members of the astacin family. Such rearrangement affects the activation domain and the new N-terminal segment, which becomes completely buried and salt-bridged to active-site residues. These interactions and a precisely trimmed N terminus are pivotal for the structural integrity and stability of astacins (31, 32).

Despite having distinct chain traces and sequences, proastacin activation shares the following properties with that of MMP zymogens: (i) the prosegments run in opposite direction to a substrate across the cleft; (ii) the catalytic zinc-bound solvent molecule is replaced by a side-chain atom of a conserved residue from the prosegment; (iii)
the new N terminus establishes a salt bridge with a residue from the mature enzyme; and (iv) activation entails initial cleavage by other proteinases, although the last step is autolytic. As for other MPs, no comparable accumulating similarities are encountered with thezymogens and activation processes of funnelsins, lysostaphins, and thermolysins. With trypsin-like serine proteinases, proastacin activation shares (i) an internal salt bridge of the new N terminus with an acidic residue directly adjacent to a main catalytic residue and (ii) an activation domain that undergoes a disorder-order transition to render a functional substrate-binding cleft. Accordingly, the present novel activation mechanism contains elements from two distinct classes of peptidases and contributes to a better understanding of enzyme maturation processes.

Acknowledgments—We acknowledge the help provided by EMBL and European Synchrotron Radiation Facility local contacts. We thank Katja Lotz, Birgit Meinke, and Barbara Prommnitz for excellent technical assistance in protein expression and purification and Robin Rycroft for helpful suggestions for the manuscript.

REFERENCES

1. Holzer, H., and Heinrich, P. C. (1980) Annu. Rev. Biochem. 49, 63–91
2. Khan, A. R., and James, M. N. (1998) Protein Sci. 7, 815–836
3. Bryan, P. N. (2002) Chem. Rev. 102, 4805–4816
4. Mittl, P. R., and Grütter, M. G. (2006) Curr. Opin. Struct. Biol. 16, 769–775
5. Vendrell, J., Querol, E., and Avilés, F. X. (2000) Biochim. Biophys. Acta 1477, 284–298
6. Gomis-Ruth, F. X. (2008) Curr. Opin. Struct. Biol. 19, 578–590
7. Bochtler, M., Odintsov, S. G., Marcyjaniak, M., and Sabala, I. (2004) Protein Sci. 13, 854–861
8. Odintsov, S. G., Sabala, I., Marcyjaniak, M., and Bochtler, M. (2004) J. Mol. Biol. 335, 775–785
9. Firczuk, M., Mucha, A., and Bochtler, M. (2005) J. Mol. Biol. 342, 578–590
10. Demidyuk, I. V., Gasanov, E. V., Safina, D. R., and Kostrov, S. V. (2008) Protein J. 27, 343–354
11. Demidyuk, I. V., Gromova, T. Y., Polyakov, K. M., Melik-Adamyan, W. R., Kuranova, I. P., and Kostrov, S. V. (2010) J. Biol. Chem. 285, 2003–2013
12. Bode, W., Gomis-Ruth, F. X., and Stöcker, W. (1993) FEBS Lett. 331, 134–140
13. Stöcker, W., Grams, F., Baumann, U., Reinecker, P., Gomis-Ruth, F. X., McKay, D. B., and Bode, W. (1995) Protein Sci. 4, 823–840
14. Gomis-Ruth, F. X. (2003) Mol. Biotechnol. 24, 157–202
15. Gomis-Ruth, F. X. (2009) J. Biol. Chem. 284, 15353–15357
16. Stöcker, W., and Bode, W. (1995) Curr. Opin. Struct. Biol. 5, 383–390
17. Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., and Tryggvason, K. (1999) Science 284, 1667–1670
18. Morgunova, E., Tuuttila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., and Tryggvason, K. (1999) Science 284, 1667–1670
19. Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., and Van Wart, H. E. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 364–368
20. Van Wart, H. E., and Birkedal-Hansen, H. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 5578–5582
21. Vallee, B. L., and Auld, D. S. (1990) Biochemistry 29, 5647–5659
22. Looechel, F., Overgaard, M. T., Oxvig, C., Albrechtsen, R., and Wewer, U. M. (1999) J. Biol. Chem. 274, 13427–13433
23. Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., and Van Wart, H. E. (1990) Proc. Natl. Acad. Sci. U.S.A. 87, 364–368
24. Vallée, B. L., and Auld, D. S. (1990) Biochemistry 29, 5647–5659
25. Looechel, F., Overgaard, M. T., Oxvig, C., Albrechtsen, R., and Wewer, U. M. (1999) J. Biol. Chem. 274, 13427–13433