Published in final edited form as:
Science. 2013 March 29; 339(6127): 1600–1604. doi:10.1126/science.1232048.

Structure of the Integral Membrane Protein CaaX Protease Ste24p

Edward E. Pryor Jr.[1,2,*], **Peter S. Horanyi**[1,2,*], **Kathleen M. Clark**[1,3,*], **Nadia Fedoriw**[1,3,*], **Sara M. Connelly**[1,3], **Mary Koszelak-Rosenblum**[1,4], **Guangyu Zhu**[1,4], **Michael G. Malkowski**[1,4,5], **Michael C. Wiener**[1,2,†], and **Mark E. Dumont**[1,3,6,†]

[1]Membrane Protein Structural Biology Consortium

[2]Department of Molecular Physiology and Biological Physics, Box 800886, University of Virginia, Charlottesville, VA 22908-0886

[3]Department of Pediatrics, P.O. Box 703, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642

[4]Hauptman-Woodward Institute, 700 Ellicott Street, Buffalo, NY 14203

[5]Department of Structural Biology, State University of New York at Buffalo, 700 Ellicott Street, Buffalo, NY 14203

[6]Department of Biochemistry and Biophysics, P.O. Box 712, University of Rochester School of Medicine and Dentistry, Rochester, NY 14642

Abstract

Post-translational lipidation provides critical modulation of the functions of some proteins. Isoprenoids (i.e., farnesyl or geranylgeranyl groups) are attached to cysteine residues in proteins containing C-terminal CaaX sequence motifs. Isoprenylation is followed by cleavage of the aaX amino acid residues and, in some cases, by additional proteolytic cuts. We determined the crystal structure of the CaaX protease Ste24p, a zinc metalloprotease catalyzing two proteolytic steps in the maturation of yeast mating pheromone **a**-factor. The Ste24p core structure is a ring of seven transmembrane helices enclosing a voluminous cavity containing the active-site and substrate binding groove. The cavity is accessible to the external milieu via gaps between splayed transmembrane helices. We hypothesize that cleavage proceeds via a processive mechanism of substrate insertion, translocation and ejection.

†To whom correspondence should be addressed. mwiener@virginia.edu (M.C.W.); mark_dumont@urmc.rochester.edu (M.E.D.).
*These authors contributed equally to this work.

Author contributions (authorship of relevant Supplementary Materials: Materials & Methods sections denoted by *): target selection & yeast homology screening (M.E.D.*); cloning (S.M.C); expression, solubilization, purification, and functional characterization with accompanying figure S3 (K.M.C.*, N.F.*, M.E.D.*); preliminary crystallization & diffraction screening (M.K.-R., G.Z., M.G.M.*); SmSte24p crystallization (E.E.P.*), SmSte24p crystallography (E.E.P.*, P.S.H.*, M.C.W.). M.C.W. conceptualized and wrote the main paper, with the following exceptions: introduction (M.E.D.), figures & captions (E.E.P, P.S.H., M.C.W.). All authors critically read the manuscript prior to submission. In particular, M.E.D. contributed significantly to the editing, revision and improvement of the manuscript.

The authors declare no competing financial interests.

Isoprenoid groups are conjugated to proteins via cysteine residues of CaaX acceptor sequences in which the cysteine attachment site is followed by two aliphatic amino acid residues and one unspecified residue at the protein C-terminus. Isoprenylation is generally accompanied by two subsequent processing steps, proteolytic cleavage of the aaX residues and carboxymethylation of the newly exposed carbonyl group of the modified cysteine residue (fig. S1). Some isoprenylated proteins also undergo additional proteolytic processing, including an additional cleavage by the same protease that initially removes the aaX residues. At least two classes of enzymes are responsible for the cleavage of isoprenylated proteins and peptides. One of these is the ras-converting enzyme (Rce) family of Type II prenyl proteases, responsible for proteolytic processing of signal-transducing proteins including Ras (1, 2) and the G_{γ} subunits of heterotrimeric G protein complexes (3). The other is the Ste24p family of Type I prenyl proteases, first identified in yeast based on its role in maturation of the mating pheromone **a**-factor (4–6). Extensive characterization of the role of Ste24p in **a**-factor processing has been conducted in the yeast system (7). The proteolytic activity of Ste24p requires zinc, consistent with the fact that Ste24p contains the zinc metalloprotease signature motif HEXXH (8, 9).

A human ortholog of Ste24p, ZMPSTE24 (**Z**inc **M**etallo**P**rotease **STE24**), can complement the full function of yeast Ste24p (6). The only known substrate for ZMPSTE24 is prelamin A, the precursor to the nuclear intermediate filament protein lamin A. Lamins provide mechanical stability to the nuclear envelope, function as scaffolds for localization of other proteins and for cytoskeletal attachment, regulate chromatin, and are implicated in transcription and DNA repair and replication (10). Mutations in either ZMPSTE24 or the processing site of prelamin A are associated with a spectrum of premature-aging diseases referred to as progeria (11). The severity of different forms of progeria is reported to be correlated with extent of loss of ZMPSTE24 activity (12). Ste24p is localized to the endoplasmic reticulum membrane. Its proteolytic activity requires zinc, consistent with the fact that Ste24p contains the zinc metalloprotease signature motif HEXXH (8, 9).

Ste24p from *Saccharomyces cerevisiae* (ScSte24p) has been overexpressed previously in *S. cerevisiae* cells and purified (9, 15). To identify forms of the protein with enhanced stability and suitability for crystallization, we cloned and purified orthologs from nine yeast species closely related to *S. cerevisiae*. *S. mikatae* Ste24p (SmSte24p) is 96% identical to ScSte24p, and is 37% identical to *H. sapiens* ZMPSTE24 (fig. S2). Purified SmSte24p is enzymatically active (fig. S3), and we obtained crystals of this protein that diffracted anisotropically to 3.1 Å resolution (and isotropically to 3.9 Å resolution). After obtaining a native dataset and proceeding to make selenomethionine-containing SmSte24p (16), we discovered that the Structural Genomics Consortium (SGC) had solved the structure of human ZMPSTE24. Due to SGC's Open Access policy, the coordinates were deposited into the protein database prior to publication (PDB:4AW6 (17)). Thus, we solved the structure of SmSte24p by a combination of molecular replacement (MR) and single-wavelength anomalous diffraction (SAD) of the bound catalytic zinc atoms.

The combination of MR and experimental SAD dispersion phases, along with application of non-crystallographic symmetry and solvent-density modification, yielded interpretable electron density maps (fig. S4). The anisotropic diffraction data likely results from the high

solvent content (~80%) of the crystal, and the marked asymmetry of crystal contacts (fig. S5). The anisotropic data between 3.9 and 3.1 Å resolution comprise a significant fraction of the entire dataset used for structure determination and refinement, and increase the total number of reflections by ~30% compared to the isotropic data. The structure was refined to R and R_{free} values of 0.270 and 0.293, respectively, with good stereochemistry (Table S1). Structure determination included extensive use of omit maps and real space correlation coefficients (figs. S4 and S6, Tables S2–S4). The asymmetric unit contains two nearly identical SmSte24p protomers (all-atom RMSD: 0.7 Å) arranged as an antiparallel dimer. The protein construct present in the crystal is the 461-residue full-length protein with eight additional residues that remain after cleavage of the C-terminal affinity tags. The refined structure of chain A consists of residues 9–106, 114–338 and 347–445; that of chain B consists of residues 11–103 and 109–446. The RMSD between SmSte24p and ZMPSTE24 is 1.2 Å.

SmSte24p is an integral membrane protein comprising seven transmembrane helices that surround a large cavity (~14,000 $Å^3$) contained nearly completely within the membrane interior (Fig. 1). The transmembrane helices range from 28 to 44 residues in length, and form a fenestrate surface with gaps of up to 10 Å diameter between helices. The cavity is sufficient to contain a 10 kDa protein or 450 water molecules. Helices III–VII are kinked or sharply curved, with helices III and IV containing proline residues at the sites of the kinks. Helix VI contains the zinc metalloprotease HEXXH motif. Two luminal loops, between helices II and III and between helices VI and VII, contain short α-helices comprised of 12 and 7 residues, respectively (fig. S7A). These short helices, the other luminal loops and the canted transmembrane helices combine to cap the luminal surface of the structure. The cytosolic-facing surface of the protein contains two additional domains: a 70-residue mixed α-helical/β-strand Loop 5 Domain (L5D) and a 44-residue α-helical C-Terminal Domain (CTD) (fig. S7A). The L5D and CTD block the large gaps between helices V and VI and between helices VII and I, respectively, and also cap the cytoplasmic surface of the structure (fig. S7B).

Although no zinc was added during the expression, purification or crystallization of SmSte24p, x-ray fluorescence scans of crystals clearly indicated a peak at the zinc edge (data not shown), and diffraction data collected at this peak yielded a single prominent 5σ anomalous difference Fourier peak in each protomer (Fig. 2A). The soluble protein thermolysin is the most thoroughly-characterized zinc metalloprotease, with a large number of available crystal structures and extensive mechanistic analyses (18–22). Comparison of the active-site of thermolysin to the zinc-binding site of SmSte24p reveals a striking structural congruence of nearly all of the residues responsible for zinc coordination, catalysis and substrate recognition (Fig. 2). In particular, an eight residue active-site substructure in thermolysin is nearly identical to the same substructure in Ste24p (Fig. 2C). Similarity between the active-sites of thermolysin and an integral membrane protein zinc metalloprotease was previously noted for the archaebacterial protein site-2 protease (S2P) (23). Ste24p and S2P have different topologies; also, unlike Ste24p, which processes peptides expected to reside at the membrane surface, S2P catalyzes intramembrane proteolysis of transmembrane proteins (24–26).

The interactions between Ste24p and its substrates that determine the specificity of the interaction remain poorly experimentally characterized; to date, only mutations immediately C-terminal to each of the cleavage sites in the **a**-factor precursor have been shown to affect cleavage efficiencies in *in vivo* studies (27, 28). To characterize a likely substrate binding site in SmSte24p, we first produced a composite model of thermolysin with an occupied active-site based on three structures of thermolysin in complex with substrates and inhibitors (Fig. 3A). We then generated an electron density envelope corresponding to the composite contents of the thermolysin active-site (Fig. 3B) and mapped this onto the structure of SmSte24p via the same transformation matrix used to calculate the RMSD between the two enzymes' active-site residues. Finally, we fitted a farnesylated peptide fragment of **a**-factor precursor (including its scissile bond) to this density (Fig. 3C,D). The modeled position of the substrate places the farnesyl attachment site at a position adjacent to the internal cavity that affords ample volume for it to adopt a variety of conformations (Fig. 3C,D). However, the modest resolution of our current structure precludes a detailed interpretation of this modeling.

The substrate-binding groove, formed by L5D and CTD, is contained at the cytoplasmic face of the internal transmembrane cavity of SmSte24p. The groove, nearly coplanar with the membrane surface and approximately 40 Å in length, traverses the cavity (Fig. 4A,B). Each end of the groove is located between one of the pairs of splayed transmembrane helices (fig. S7B). The high degree of similarity between the active-sites of SmSte24p and thermolysin strongly suggests that the polarity (N-to-C direction) of peptide/protein substrates with respect to the active-site is identical in the two enzymes. Thus, SmSte24p substrates can be expected to be oriented in the groove with their C-termini pointing toward the portal between helices VII and I of Ste24p and their N-termini pointing towards the portal between helices V and VI (Fig. 4C). The localization of residues involved in enzymatic activity to the cytosolic-proximal portion of the cavity is reflected in higher sequence identity and conservation between SmSte24p and ZMPSTE24 in this part of the cavity (fig. S8).

These observations, together with previously-reported functional similarity between Ste24p and ZMPSTE24, particularly the ability of ZMPSTE24 to fully complement a deletion of *STE24* in processing yeast **a**-factor (6), suggest a possible pathway for substrate processing by CaaX proteases. In considering a small substrate such as the **a**-factor precursor (36 residues in length), substrate binding could occur by entry of either its C- or N-terminus into the binding groove. However, prelamin A contains a folded domain of 647 residues N-terminal to both ZMPSTE24 cleavage sites (fig. S1A). Since it is unlikely that such a large domain could gain access to the active-site, the most likely pathway for entry of this substrate (and, by inference, all the substrates of this type of enzyme) into the active-site is via insertion of its C-terminus (Fig. 4D). The second proteolytic cleavages of **a**-factor precursor and prelamin A could occur via processive further extension of the substrate C-terminus into the substrate groove or cavity following the first cleavage (Fig. 4E). However, the presence of the large folded domain of ZMPSTE24 dictates that withdrawal of the final N-terminal cleavage products would have to occur via reversal of the initial route of entry. Since L5D and CTD block the gaps between helices V & VI and helices VII & I, respectively (fig. S7B), movement of one or both domains may be required for substrate

processing. However, relatively modest conformational changes or movement may suffice to enlarge the fenestrations that already exist in these structural regions (Fig. 3D). Interestingly, most of the missense mutations in human ZMPSTE24 that result in progeria map to the putative entrance and exit portals (fig. S9).

A possible role for the voluminous internal cavity is suggested by the consumption of a water molecule in each turnover of proteolytic enzymes. While biological membranes are quite water-permeable, the dwell-time of water in the membrane interior is low and maximizing the effective concentration of water may promote cleavage reaction by mass action. Thus, the cavity may serve as a water reservoir to facilitate CaaX protease substrate processing. Other integral membrane proteases appear to rely upon different mechanisms for obtaining access to water. The catalytic serine of rhomboid proteases is positioned at the base of a funnel-shaped periplasmic-accessible cavity that may be gated (29). The zinc atom of S2P is accessible to the cytoplasm via a narrow water-permeable channel present in the closed (to substrate) conformation of the enzyme (23). The role of the large cavity, the mechanism of specific recognition of cleavage sites with divergent sequences, the unusual dual cleavage process, and the possible role of the farnesyl group in recognition are some of the questions raised by this provocative structure.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health PSI:Biology grant U54 GM094611 for membrane protein structural genomics. Data were collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. We thank Drs. Patrick Loll, Robert Nakamoto, and Jochen Zimmer for useful discussion. Atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession number 4IL3.

References and Notes

1. Dolence JM, Steward LE, Dolence EK, Wong DH, Poulter CD. Biochemistry. 2000; 39:4096. [PubMed: 10747800]
2. Kim E, et al. J Biol Chem. 1999; 274:8383. [PubMed: 10085069]
3. Otto JC, Kim E, Young SG, Casey PJ. J. Biol. Chem. 1999; 274:8379. [PubMed: 10085068]
4. Boyartchuk VL, Ashby MN, Rine J. Science. 1997; 275:1796. [PubMed: 9065405]
5. Fujimura-Kamada K, Nouvet FJ, Michaelis S. J Cell Biol. 1997; 136:271. [PubMed: 9015299]
6. Tam A, et al. J Cell Biol. 1998; 142:635. [PubMed: 9700155]
7. Michaelis S, Barrowman J. Microbiol Mol Biol Rev. 2012; 76:626. [PubMed: 22933563]
8. Jongeneel CV, Bouvier J, Bairoch A. FEBS Lett. 1989; 242:211. [PubMed: 2914602]
9. Tam A, Schmidt WK, Michaelis S. J Biol Chem. 2001; 276:46798. [PubMed: 11581258]
10. Dittmer TA, Misteli T. Genome Biol. 2011; 12:222. [PubMed: 21639948]
11. Worman HJ. J Pathol. 2012; 226:316. [PubMed: 21953297]
12. Barrowman J, Wiley PA, Hudon-Miller SE, Hrycyna CA, Michaelis S. Hum Mol Genet. 2012; 21:4084. [PubMed: 22718200]
13. Coffinier C, et al. Proc Natl Acad Sci U S A. 2007; 104:13432. [PubMed: 17652517]

14. Hudon SE, et al. Biochem Biophys Res Commun. 2008; 374:365. [PubMed: 18639527]
15. Clark KM, et al. Protein Expr Purif. 2010; 71:207. [PubMed: 20045057]
16. Malkowski MG, et al. Proc Natl Acad Sci U S A. 2007; 104:6678. [PubMed: 17426150]
17. Quigley A, et al. Science. 2013; 339:1604. [PubMed: 23539603]
18. Hangauer DG, Monzingo AF, Matthews BW. Biochemistry. 1984; 23:5730. [PubMed: 6525336]
19. Kester WR, Matthews BW. Biochemistry. 1977; 16:2506. [PubMed: 861218]
20. Matthews BW. Acc Chem Res. 1988; 21:333.
21. Matthews BW, et al. Nat New Biol. 1972; 238:41. [PubMed: 18663850]
22. Matthews BW, Jansonius JN, Colman PM, Schoenborn BP, Dupourque D. Nat New Biol. 1972; 238:37. [PubMed: 18663849]
23. Feng L, et al. Science. 2007; 318:1608. [PubMed: 18063795]
24. Brown MS, Ye J, Rawson RB, Goldstein JL. Cell. 2000; 100:391. [PubMed: 10693756]
25. Golde TE, Eckman CB. Sci STKE. 2003; 2003:RE4. [PubMed: 12621149]
26. Wolfe MS, Kopan R. Science. 2004; 305:1119. [PubMed: 15326347]
27. Huyer G, et al. Eukaryot Cell. 2006; 5:1560. [PubMed: 16963638]
28. Trueblood CE, et al. Mol Cell Biol. 2000; 20:4381. [PubMed: 10825201]
29. Urban S. Biochem J. 2010; 425:501. [PubMed: 20070259]
30. Dundas J, et al. Nucleic Acids Res. 2006; 34:W116. [PubMed: 16844972]
31. Ho BK, Gruswitz F. BMC Struc Biol. 2008; 8:49.
32. Schmidt WK, Tam A, Michaelis S. J Biol Chem. 2000; 275:6227. [PubMed: 10692417]
33. Klauda JB, et al. J Phys Chem B. 2010; 114:7830. [PubMed: 20496934]
34. Berka K, et al. Nucleic Acids Res. 2012; 40(Web server issue):W222. [PubMed: 22553366]

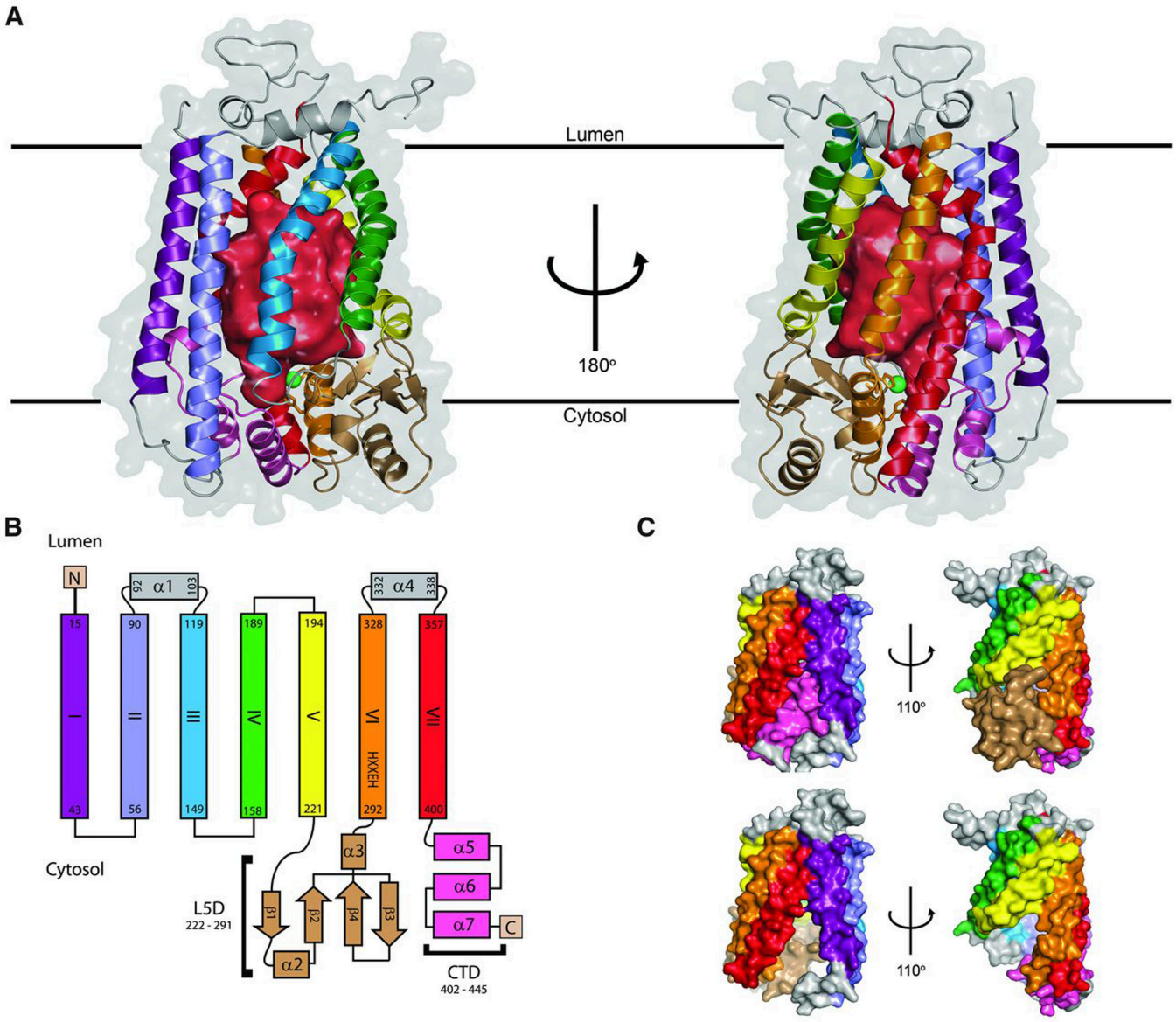

Fig. 1.

Overall structure of SmSte24p. Ribbon diagram of Ste24p highlighting pronounced kinks in transmembrane helices III–VII. The red surface in the center of the protein represents the boundary of the ~14,000 $Å^3$ interior cavity, calculated with CASTp (30), visualized and rendered with HOLLOW (31). The green sphere is the catalytic zinc atom.

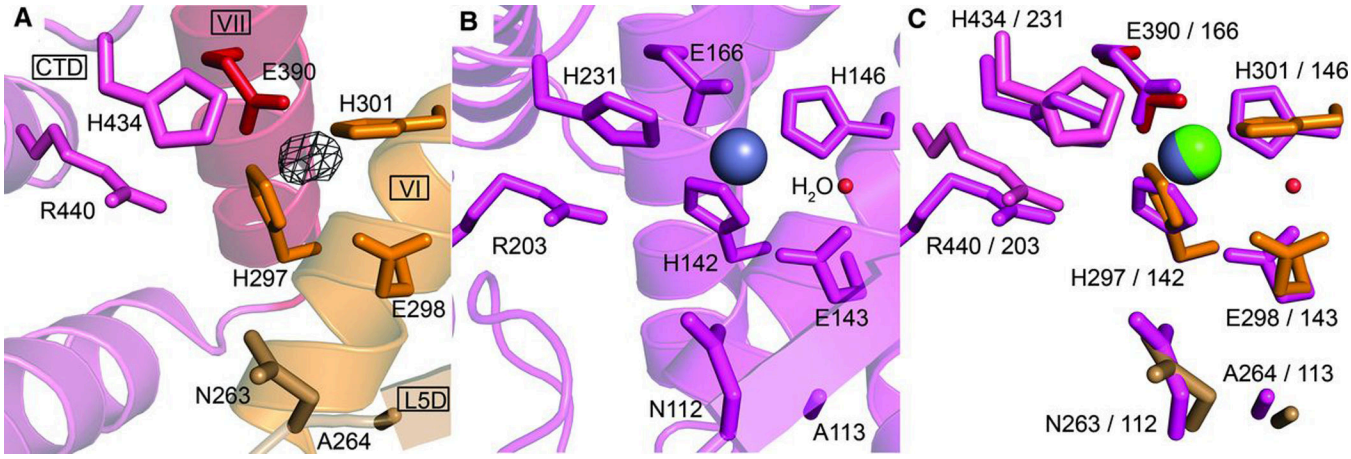

Fig 2.

Analysis of the SmSte24p active-site. **(A)** The active-site of SmSte24p contains the HEXXH-motif in helix VI (H297–H301), the catalytic residue (E390), and substrate coordinating residues N263, A264, H434, and R440. Mutation of H297, E298 or H301 diminishes Ste24p function (4, 32). The catalytic zinc atom is represented by anomalous electron difference density contoured at 5σ. **(B)** The active-site of thermolysin (PDB:1LND) showing the HEXXH-motif (H142–H146), the catalytic residue (E166) and the substrate coordinating residues N112, A113, R203 and H231. **(C)** Overlay of thermolysin and SmSte24p active-sites showing their similarity (RMSD of overlaid residues: 0.5 Å).

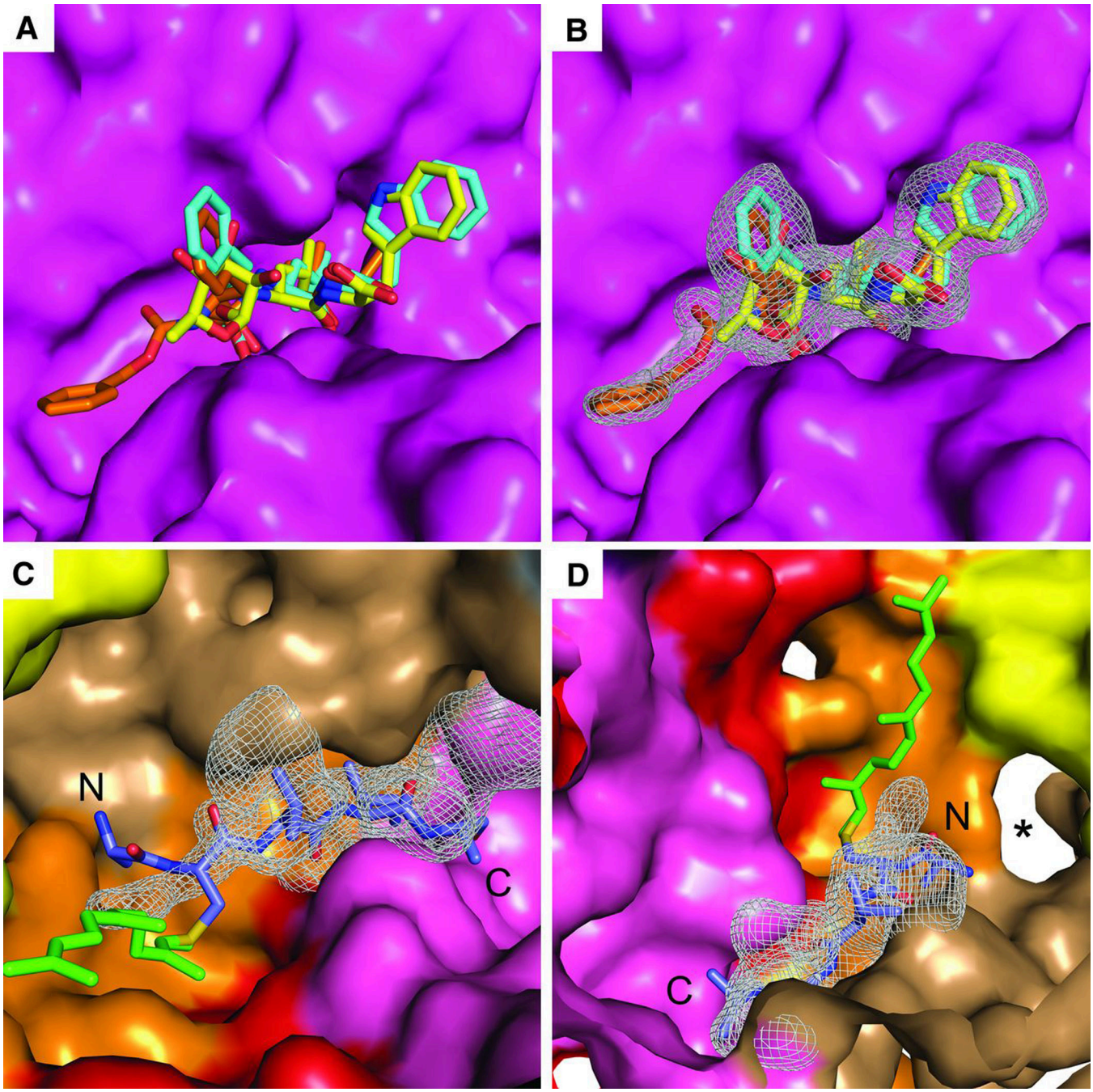

Fig. 3.

Modeling of pre-**a**-factor binding to SmSte24. **(A)** Surface representation of the substrate binding pocket of thermolysin with three different substrate/inhibitor molecules overlaid. **(B)** Electron density envelope calculated from (A). **(C)** Electron density envelope from (B) mapped onto the SmSte24p structure with an **a**-factor precursor peptide fragment (ACVIA) positioned within the density. **(D)** Alternate view of the modeled **a**-factor precursor peptide with attached farnesyl chain. The white area indicated by the asterisk identifies fenestration between helices V and VI.

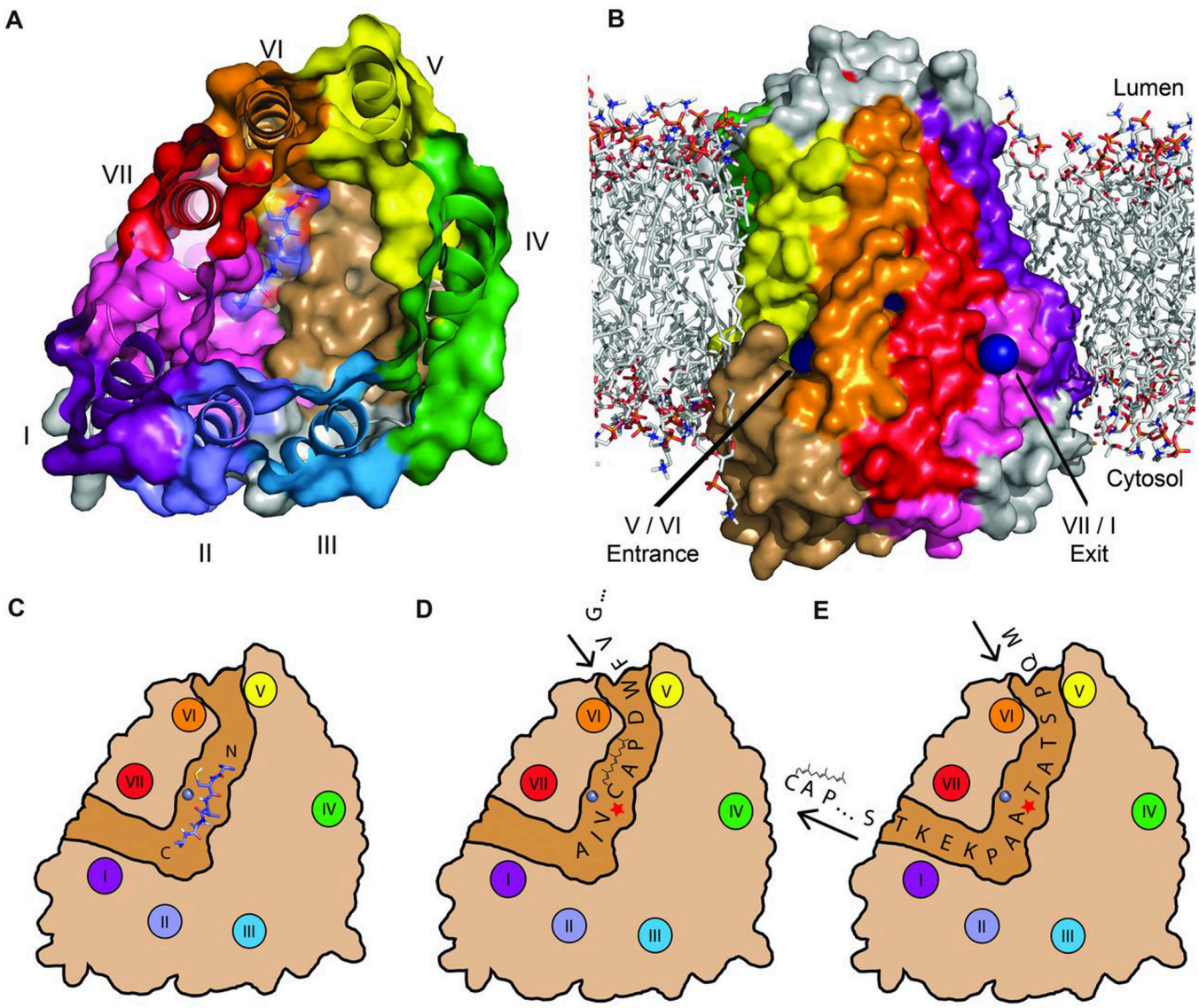

Fig. 4.

Proposed Ste24p mechanism. **(A)** Cutaway view of the SmSte24p active-site from the luminal surface. The modeled **a**-factor precursor peptide (from Fig. 3, C and D) is shown in blue. **(B)** Surface representation of SmSte24p, positioned into a lipid bilayer model obtained from a published molecular dynamics simulation (33). Blue spheres denote gaps between helices V/VI and between helices VII/I that lead to the substrate-binding groove and internal active-site contained within the internal transmembrane cavity (Fig 1A). **(C)** Schematic representation, to scale, of the continuous groove that traverses the cavity (determined with MOLE 2.0 (34)). The ~40 Å tunnel can accommodate ~13 residues of substrate. **(D)** Insertion of the **a**-factor precursor C terminus into the entrance portal formed by helices V and VI. Cleavage by SmSte24p occurs at the CaaX motif indicated by the red star. **(E)** Subsequent substrate translocation, egress through the VII/I portal, and positioning of the second cleavage site.