Conformational diversity within unique amino acid sequences is observed in diseases like scrapie and Alzheimer’s disease. The molecular basis of such diversity is unknown. Similar phenomena occur in subtilisin, a serine protease homologous with eukaryotic pro-hormone convertases. The subtilisin propeptide functions as an intramolecular chaperone (IMC) that imparts steric information during folding but is not required for enzymatic activity. Point mutations within IMCs alter folding, resulting in structural conformers that specifically interact with their cognate IMCs in a process termed “protein memory.” Here, we show a mechanism that mediates conformational diversity in subtilisin. During maturation, while the IMC is autocleaved and subsequently degraded by the active site of subtilisin, enzymatic properties of this site differ significantly before and after cleavage. Although subtilisin folded by Ile<sup>48</sup> → Thr the IMC (IMC<sup>1-48T</sup>) acquires an “altered” enzymatically active conformation (Sub<sup>1-48T</sup>48) significantly different from wild-type subtilisin (Sub<sup>WT</sup>), both precursors undergo autocleavage at similar rates. IMC cleavage initiates conformational changes during which the IMC continues its chaperoning function subsequent to its cleavage from subtilisin. Structural imprinting resulting in conformational diversity originates during this reorganization stage and is a late folding event catalyzed by autocleavage of the IMC.

Subtilisins constitute a family of serine proteases that have served as model systems for understanding the structural origin of protease function and specificity (1). Cloning of the gene for subtilisin E from Bacillus subtilis revealed that it is synthesized with an N-terminal precursor domain that is referred to as the propeptide (2). Although several biochemical studies on subtilisin have indicated that the 77-mer propeptide was not part of the 275-amino acid active protease (3, 4), Escherichia coli expression experiments revealed that the propeptide of subtilisin was essential for proper activation and secretion of the protease in vivo (5–7). Soon thereafter, the propeptide of α-lytic protease (a serine protease not evolutionarily related to subtilisin) was also shown to be essential for proper activation of its protease domain (8). Although these results were surprising at the time, propeptides are now known to be required for production of a large number of proteases. Propeptide-dependent systems have since then been identified both in prokaryotes and eukaryotes (9). Virtually all known extracellular bacterial proteases have propeptides (10). There are now examples of serine (5, 11), aspartyl (12), cysteine (13), and metalloproteases (14) synthesized as precursor pro-proteases. Propeptides seem to perform two distinct functions: folding and inhibition of their protease domains (15). In doing so they act as single turnover enzymes due to their tight binding with folded proteases and their proteolytic sensitivity. As propeptides facilitate correct folding they were classified as molecular chaperones (16), but due to differences that exist between molecular chaperones, propeptides were also termed “intramolecular chaperones” (17). Although many proteases require their propeptides to facilitate their folding, these findings do not constitute a rule. In addition to proteases, proteins such as growth factors, neuropeptides, hormones, and plasma proteins also require their propeptides for correct folding (18–21). Studies on the conformations of propeptides of subtilisin (22–24) and carboxypeptidase Y (25) have shown them to be randomly structured. In the vicinity of the prefolded protease domain, the propeptide of subtilisin adopts an α-β conformation (22–24).

Anfinsen demonstrated that all the information necessary for folding of a protein into an active conformation resides in the amino acid sequence of that protein (26). Based on Anfinsen’s observation, the native state of a protein is generally believed to be the global free energy minimum (27). However, there is increasing evidence that kinetically selected states play a role in the biological function of some proteins (28). For example, the serpin plasminogen activator inhibitor 1 folds into an active structure and then converts slowly to a more stable, but low activity “latent” conformation (29). Thus, the folding of plasminogen activator inhibitor 1 is apparently under kinetic control. When α-lytic protease (11, 30) and subtilisin (31, 32) are folded in the absence of their propeptides, they fold into partially structured molten globule intermediates. These stable but inactive intermediates can convert into active conformations upon addition of their propeptides (30, 31) and suggest that propeptides promote folding of their proteases by direct stabilization of the rate-limiting folding transition state (30, 31). Propeptides seem to be essential only during late stages of the folding process because they help overcome a kinetic block in the folding pathway. A similar state is formed even in the presence of the propeptide (33), and autoprocessing may occur in this intermediate state.

Point mutations within the propeptides can affect their folding function (34, 35). As a result, identical protein sequences can display altered, enzymatically active conformations with specificity toward the propeptide that mediates folding (36). This finding helped introduce the concept termed “protein memory” and suggests that propeptide domains function as steric chaperones (28, 36). Altered protein folding also occurs in prion diseases wherein the protein can exist in two different conformations with invariant amino acid sequences, although...
this folding does not require propeptides (37). Therefore, pro-
proteins can display conformationally diverse biologically active
states, and the native state of proteins may not necessarily
reside at a thermodynamic minimum but may be ones that are
most easily accessible (28, 36).

In the present manuscript we describe a mechanism that
mediates conformational diversity in subtilisin. Our results
show that the propeptide continues its chaperoning function
subsequent to its cleavage from subtilisin and that structural
imprinting that results in conformational diversity originates
during this reorganization stage.

MATERIALS AND METHODS

Site-directed Mutagenesis—The desired mutations within the prope-
tide will be introduced through site-directed mutagenesis using polymer-
ase chain reaction (38). After site-directed mutagenesis the plasmids were
sequenced using an automated sequencer (Applied Biosystems, Inc.,
model 310).

Protein Expression and Purification—The plasmid (pET11a) carrying
mutant and wild-type proteins were expressed in E. coli strain BL21(DE)
that is grown on M9 medium. Propeptides were isolated from inclusion
bodies by solubilization in 15–25 ml of 6 M guanidine HCl (39). After
overnight incubation at 4 °C, the insoluble materials were removed
through centrifugation. The supernatants were then dialeyzed against
50 mM sodium phosphate buffer at pH 5.0. Purification was carried out
using a cation-exchange column (CM-Sephadex-50) with a linear
gradient of NaCl (0–0.4 M) followed by high pressure liquid chromatog-
raphy (C18-reverse phase) with a linear acetonitrile gradient (0–65%).

Because the propeptide is also a substrate for the reaction, measure-
ment of the inhibition constants is more complex. Therefore, the inhibi-
tion experiments were carried out under pseudo-first-order kinetics
where the lowest propeptide concentrations were greater than 10-fold
excess. Enzyme concentrations were set at suitably low levels to give
measurable rates of substrate hydrolysis, with observable state of in-
hibitor binding over the steady state time scale. The propeptide
concentrations vary between 1 and 0.1 M. Initial substrate concentration
was 3 mM, and less than 10% of the substrate was degraded throughout
the experiment. Reactions were initiated by addition of SubWT or SubI-
wt. After thorough mixing, substrate cleavage was monitored by re-
cording at 415 nm the release of p-nitroaniline through changes
in absorbance at 405 nm. Reading were collected at 15 s using a Bio-Rad
UV microplate reader, and the velocity of reaction was estimated (36).

Circular Dichroism Studies—CD measurements are performed on an
automated AVIV 60DS spectrophotometer fitted with a thermostated
cell holder that is controlled by an on-line temperature control unit.
Quartz rectangular cells (Precision Cells, Hicksville, NY) with a path
length of 1 mm are used. For CD studies, modified subtilisin in 50 mM
Tris-HCl, pH 7.0, containing 0.5 M (NH4)2SO4 will be used. Solutions are
filtered through a 0.22-μm filter before measurement. Scans were carried out at wavelengths between 260 and 190 nm in a cuvette with a 1-mm path length maintained at 10 °C. In the
thermal unfolding experiments, temperatures were increased from 10
to 90 °C in 1 °C intervals, with a 30-s equilibration at each temperature.
Data were collected at each temperature for 10 s. Protein solutions
contain 1 mM phenylmethylsulfonyl fluoride to prevent autolysis during
this procedure (36).

The in vitro maturation of the 352-residue IMC-subtilisin E
involves folding, autocleavage, and degradation of its IMC
domain (77-residues) to give mature subtilisin (275-residue). The
folding of pro-subtilisin occurs through a molten globule like
intermediate that becomes compact during or after the auto-
cleavage of the propeptide domain.

Refolding of Denatured Pro-subtilisin—Denatured precu-
sors are refolded through a step-wise dialysis procedure that
gradually reduces urea concentration (see “Materials and
Methods”). Fig. 1a depicts CD spectra of denatured IMCwt.
S221C-subtilisin. Extensive dialysis (48 h) against buffer that contains 1.5 M urea, pH 7.0, induces significant secondary structure within the uncleaved precursor (Fig. 1a). This precursor binds with ANS (Fig. 1b), a hydrophobic dye that fluoresces upon interacting with hydrophobic patches within proteins. Complete removal of urea through dialysis induces the subtilisin domain to cleave its IMC. The resulting autocleaved complex (Fig. 1c) was purified using gel filtration and displays a secondary structure pattern similar to native subtilisin and fluoresces approximately 8-fold less than the uncleaved precursor (Fig. 1b). Differences in ANS fluorescence intensities suggest that a structural reorganization that alters solvent accessible hydrophobic surfaces within the protein coincides with autocleavage.

Characterization of Autoprocessing and Degradation Activity of Pro-Ser221Cys-subtilisin—A Ser221 → Cys substitution at the active site of subtilisin blocks maturation after autocleavage, resulting in a stable stoichiometric IMC-S221C-subtilisin complex (43), whose x-ray structure was recently solved at 2 Å resolution (44, 45). Although active site residues Asp32, His64, and Ser221 within subtilisin mediate both autocleavage and degradation activities are different from each other. Mapping the Effect of the Propeptide Mutation on the Maturation Pathway—IMCWT-S221C-subtilisin affects the folding process, and although the precursor undergoes maturation, the resulting conformation of the protease domain (SubI-48V) differs from that of SubWT. Because autocleavage and degradation represent two activities of the same active site, effects of IMC mutations on these activities may help elucidate the molecular basis of altered folding. SubI-48V retains significant enzymatic activity, and hence the mutation Ile28 → Thr was isolated in an attempt to further reduce proteolytic activity. Crystallographic studies of the IMC-subtilisin complex (44, 45) shows that Ile28 is surrounded by hydrophobic side chains (Fig. 2a). Thromine, a polar residue, may disrupt the hydrophobic core and augment the phenomenon of altered folding. Upon renaturation, IMCWT-S221C-subtilisin cleaves its IMC domain (Fig. 2b) to form a complex that is further purified using gel filtration. Similarly, IMC48T-S221C-subtilisin undergoes maturation forming mature SubI-48T (Fig. 2c) with a well defined secondary structure (Fig. 2e). The purified autocleaved complexes IMCWT and IMC48T-S221C-subtilisin interact differently with ANS, with IMC48T-S221C-subtilisin exhibiting a 4-fold greater fluorescence intensity than IMCWT-S221C-subtilisin complex. Similarly, SubI-48T fluoresces with an magnitude approximately 2-fold of that of SubWT (Fig. 2d). These results suggest that subtilisin domains folded by the IMC48T display exposed hydrophobic surfaces that are absent in domains folded using IMCWT. Far ultraviolet CD spectra indicate that SubI-48T and SubWT are well folded and have similar but not identical conformations (Fig. 2c). The pH profile of the SubI-48T is depicted in Fig. 1d, and it displays maximum proteolytic activity at approximately pH 8.2, whereas autocleavage is maximum at pH 7.0.

Thermostability of Autoprocessed Complexes—Thermosta-
bilities of mature protease domains Sub\(^{WT}\) and Sub\(^{I-48T}\) and the corresponding autocleaved complexes were determined by monitoring changes in negative ellipticity at 222 nm (Fig. 3a). The autoprocessed complexes used in these experiments were separated from the unautoprocessed precursors using gel filtration chromatography as described earlier (41). Sub\(^{I-48T}\) (\(T_m 49.5^\circ C\)) was found to be substantially more unstable than Sub\(^{WT}\) (\(T_m 58.5^\circ C\)), whereas the IMCI-48T-S221C-subtilisin autocleaved complex (\(T_m 46.8^\circ C\)) was also more unstable than the IMC\(^{WT}\)-S221C-subtilisin complex (\(T_m 53.2^\circ C\)). Both protease domains complexed with the IMC domains melt at lower temperatures than their corresponding protease domains. This suggests that the presence of IMC domains destabilize the corresponding complexes and conversely that the degradation of the IMC domains enhances stability of cognate protease domains. It is important to note that phenylmethylsulfonyl fluoride covalently interacts with the active site Ser\(^{221}\) residue. Crystallographic results have shown that such covalent linkage may cause local perturbations but does not seem to affect the rest of the backbone. The introduction of a bulky residue into the active site may therefore slightly destabilize the mature protein, causing it to lower the \(T_m\) of the mature protease. Therefore the difference between the \(T_m\) of subtilisin and the complex with its propeptide may actually be greater.

Characterization of the Binding of Propeptide with Mature Subtilisin—Fig. 3b describes the IMCI-48T and IMC\(^{WT}\) binding with Sub\(^{WT}\) and Sub\(^{I-48T}\). From these curves it is evident that IMCI-48T binds to Sub\(^{I-48T}\) more strongly (and is degraded more quickly) than with Sub\(^{WT}\), suggesting more specific recognition of Sub\(^{I-48T}\) than Sub\(^{WT}\) by IMCI-48T. Rates of autocleavage of the IMCI-48T and IMC\(^{WT}\)-S221C-subtilisin precursors are estimated as described under “Materials and Methods.” Relative rates of autocleavage of the two precursor proteins appear similar (Fig. 3c), whereas proteolytic activities of Sub\(^{WT}\) and Sub\(^{I-48T}\) toward a peptide substrate Nsuccinyl-A-A-P-F-p-nitroanilide differ significantly (Fig. 3d). Enzymatic activities of Sub\(^{WT}\) and Sub\(^{I-48T}\) were estimated as described earlier (12, 26), and Sub\(^{I-48T}\) displays a \(K_m\) of \(0.24 \times 10^{-3}\) M, whereas Sub\(^{WT}\) displays a \(K_m\) of \(2.0 \times 10^{-3}\) M. Because autocleavage occurs at similar rates whereas the proteolytic specific activity of Sub\(^{I-48T}\) is only 9.0% of that of Sub\(^{WT}\), the results indicate that Ile\(^{248}\)→Thr substitution seems to selectively affect proteolytic activity (final conformation) but not autocleavage of the maturation intermediate.

**DISCUSSION**

Based on the above results a mechanism that leads to altered folding is proposed (Fig. 4). It is unclear how the IMC initiates this folding process, but it has been speculated that two helices between residues 100 and 144 within the protease domain may be involved (44, 45). The IMC domain facilitates the precursors adopt an intermediate state that is capable of autocleavage, a process mediated by the active site of subtilisin (5–8). The properties of this active site differ from that of the mature protease because (i) it displays an optimum pH that is 1.5 units less than the mature protein; (ii) IMC-S221C-subtilisin can only autocleave its IMC but cannot degrade it to release mature S221C-subtilisin; and (iii) both IMCI-48T and IMC\(^{WT}\) uncleaved precursors bind strongly with ANS, whereas the autocleaved complexes do not. This is consistent with structural reorgan-
zation that occurs after proteolytic cleavage (33). IMC\textsuperscript{WT} and IMC\textsuperscript{WT}-S221C-subtilisin both cleave their IMC domains at similar rates, indicating that the Ile\textsuperscript{48}→Thr substitution has little effect on autocleavage. Because Sub\textsuperscript{WT} and Sub\textsuperscript{I-48T} differ in their enzymatic activity after cleavage and subsequent degradation, we conclude that altered folding and protein memory that result from IMC-mutations occurs after autocleavage, after which the IMC continues to functions as a chaperone. Therefore, autocleavage does not imply that the folding process has been completed and may represent the transition state of the folding reaction. Autocleavage and degradation are closely coupled in wild-type subtilisin, and occurrence of these two processes renders the folding process irreversible.

Although proteolytic activity of subtilisin at pH 8.5 is approximately 2-fold greater than pH 7.0 (Fig. 1d), incubation of the IMC\textsuperscript{WT}-S221C-subtilisin complex at pH 8.5 does not facilitate degradation of the IMC domain (data not shown). Changes in the optimum pH alone cannot explain the absence of proteolytic activity toward the noncovalently bound IMC domain. The protease domain in the autocleaved complex is almost identical
to the wild-type protease domain with the exception of the active site Ser\(^{221}\) → Cys substitution that causes a 10,000-fold drop in the proteolytic activity. Because uncleaved precursors display exposed hydrophobic surfaces relative to cleaved complexes, it is reasonable to speculate that distribution of charge on cleaved and uncleaved protein surfaces are different. It is known that altering surface charges through mutagenesis produces subtilisin with altered specificities and pH activity profiles with enhanced catalytic activities (48, 49). Therefore, the active site within uncleaved precursor may also display enhanced autocleavage activity. After cleavage, the resulting conformational changes alter the catalytic properties of this active site. Because Ile\(^{248}\) Thr substitution within the IMC does not affect cleavage but alters enzymatic activity of the protease domain, our results indicate that conformational diversity that leads to protein memory occurs late during the folding pathway and the IMC functions even after its autocleavage.

A sequence alignment of propeptides from the subtilisin family shows an interesting observation. Although the overall sequence identity is low, two regions dispersed over the N and C termini display significant sequence conservation (Fig. 5). These regions designated as motifs N1 and N2 contain the hydrophobic core residues (Val\(^{12}\), Phe\(^{14}\), Ile\(^{30}\), Val\(^{37}\), Leu\(^{51}\), Val\(^{56}\), Leu\(^{59}\), Val\(^{65}\), and Val\(^{68}\)) within the subtilisin propeptide, apart from Ile\(^{30}\). We speculate that motifs N1 and N2 may be critical for nucleation of the folding process, whereas the nonconserved segments between N1 and N2 may be crucial for specific interactions with their cognate protease domains, during the propagation of folding and subsequent inhibition of activity. This speculation is supported by four of our findings: (i) A substitution at position Ile\(^{30}\) by a Val allows the propeptide to exert its chaperoning function but alters the final conformation of the protein (36). This Ile\(^{30}\) residue is located in the nonconserved region flanked by motifs N1 and N2 (note that Ile\(^{248}\) was previously termed Ile\(^{248}\)). (ii) By using an Ile\(^{30}\) Thr substitution, we have shown that the structural imprinting, which results in protein memory, occurs late in the folding pathway, after the propeptide has been cleaved from the protease domain through autocatalysis. Therefore, a substitution at position 30 within the propeptide does not affect folding nucleation but dramatically alters the final conformation of the protein. (iii) Results of a random mutagenesis procedure that helped identify mutational hot spots within the propeptide show that substitutions that blocked folding of subtilisin are localized within motifs N1 and N2 (34). (iv) NMR spectroscopy on the propeptide of subtilisin E shows that residues located within motifs N1 and N2 display residue-specific conformational rigidity, a criteria that classifies a polypeptide segment as a potential nucleation site in a protein folding reaction.

Acknowledgments—We thank Drs. S. Phadatare and K. Madura for comments and suggestions, J. Liu for technical help, and Yuyun Li for cloning IMC\(^{221}\)-subtilisin in pET11a.

REFERENCES
1. Wells, J. A., and Estell, D. A. (1988) Trends Biol. Sci 13, 291–297
2. Wong, S. L., and Doi, R. H. (1986) J. Biol. Chem. 261, 10176–10181

Fig. 5. Sequence alignments of various IMC domains showing conserved motifs N1 and N2 flanking a variable region. Bold letters represent conserved residues.
