Understanding mechanisms of pH sensing in the Proprotein Convertases

---

A Thesis

Presented to

The Division of Biochemistry and Molecular Biology

Oregon Health & Science University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

Danielle Marie Williamson

August 2015

Preface

The woods are lovely, dark, and deep,
But I have promises to keep,
And miles to go before I sleep,
And miles to go before I sleep.

R. Frost

Table of Contents

List of Tables

List of Figures

Abbreviations

CD	Circular dichroism
DCSG	Dense core secretory granules
ER	Endoplasmic Reticulum
ESP	Extracellular serine protease
IMC	Intramolecular Chaperone
ISP	Intracellular serine protease
mAb	Monoclonal antibody
MAT	Mature protease domain
MD	Molecular Dynamics
MSA	Multiple sequence alignment
PC	Proprotein Convertase
PC1/3	Proprotein Convertase 1/3
PCSK	Proprotein Convertase Subtilisin-Kexin
POMC	Proopiomelanocortin
PRO	Propeptide
TGN	Trans-Golgi Network
WT	Wild type

Acknowledgments

The last four years have been a small but significant part of my academic journey. While I am not yet at the end of this adventure, there are many people who deserve recognition and thanks for their role in getting me to this milestone.

I owe a debt of gratitude to all of the individuals who have mentored and guided me thus far. First, I would like to thank Dr. Ujwal Shinde for the support, direction, and innumerable cups of coffee he has provided over the years; even before I was his student, his door was always open, and his advice usually free. I would also like to thank the members of my advisory committee, Dr David Farrens, Dr. Caroline Enns, and Dr. Larry David for their support and dedication to my training. I am also grateful to the teachers, professors and mentors who gave me the knowledge and tools to set out on this journey. Particular thanks are due to several members of the Ullman lab; Dr. Buddy Ullman, Dr. Nicola Carter, and Dr. Phil Yates, for introducing me to the world of research and parasitology, teaching me to do my first western blots and maxi preps, and their role in convincing me that this was a worthwhile endeavor. It has been ten years since I joined the lab as a summer student, and like many members of the Ullman lab, I've never really left.

Like all good adventures, the time since starting graduate school has been full of twists and turns, all of which I am glad to have been able to share with friends new and old. I have found myself surrounded by an amazing cohort of

fellow students who I look forward to having as colleagues in the future. I would especially like to thank Johannes for being a great teammate and companion during our mutual time in the lab, and Jessica, who has become so much more than a classmate since the first day she extended her foot in friendship five years ago. Additionally, I would like to thank everyone who has joined me on my adventures out of doors; I can't count how many miles of road and trail I've ran, but I am glad to have shared many of them with Ben, Greg, and Derek. I would also like to thank Jan for taking me up, down, and around several mountains, all of which has been accomplished without anyone falling off. Finally, I would like to thank Caroline and Michael, who have come to be two of my dearest friends and partners in crime here in Portland – the nights of good food, wine, and conversation have kept me fueled in body and mind.

Finally, and perhaps most importantly, words cannot fully express how grateful I am for the continued love and support of my family. As I embark upon the next phase of this adventure, nothing gives me more comfort that knowing that they are always there when I need them.

Abstract

The Proprotein Convertases are members of the subtilase superfamily of serine endoproteases, and are responsible for the processing and activation of myriad secreted proteins and peptides within the cell. The necessity for precise spatiotemporal cleavage of substrates mandates the activity of the proprotein convertases be likewise stringently controlled; like all subtilases, the propeptide plays a critical role in the compartment-specific folding, trafficking, and activation of the Proprotein Convertases. While a conserved pH-sensing histidine has previously been identified in the propeptide of furin, an understanding of the precise mechanism by which this residue is able to sense pH and affect activation of its cognate catalytic domain was unknown. Additionally, different PC paralogues are activated in different compartments of the secretory pathway, an observation that suggested there must be other, yet undescribed determinants of the pH-dependent activation of these proteases.

In the following dissertation, I present my contribution to our understanding of how the Proprotein Convertases have evolved regulate their activation via their propeptides. I have demonstrated that the propeptides contain sufficient information to dictate pH-dependent activation, and that protonation of the conserved histidine drives local unfolding to exposes the cleavage site to proteolysis *in trans*. Furthermore, I present evidence that the spatial juxtapositioning of additional histidines can “tune” the pH sensitivity of the pH-sensing histidine, thus allowing for individual Proprotein Convertases to be activated within unique compartments of the secretory pathway. Finally, I demonstrate that polymorphisms within the propeptide can impart structural and functional changes upon catalytic domains that are identical in primary sequence. These findings shed light on the structure-function relationship between propeptide and protease domains of the Proprotein Convertases, and have implications for development of therapeutics for the treatment of pathologic processes driven by aberrant Proprotein Convertase activity.

Chapter I: Propeptides, proton gradients and the Proprotein Convertases

The Prohormone Theory

Two discoveries in 1967 revealed the now fundamental cellular paradigm that active peptide hormones are derived from the endoproteolytic cleavage of larger, inactive precursor polypeptides. Michel Chrétien established that the pituitary hormone β-lipotrophin was processed to two smaller proteins, γ-lipotropic and β-melanotropin, with distinct biological functions, while independently, Don Steiner elucidated the processing of proinsulin to its active form [1]. Initially thought to only apply to a few specific hormones, the prohormone theory was expanded to many protein precursors in the following decades, including neuropeptides, growth and transcription factors, receptors, extracellular matrix proteins, bacterial toxins, and viral glycoproteins. In driving the model of precursor protein processing, these discoveries necessitated the elucidation of the complex series of post-translational modifications that defined the active state in contrast to the inactive state of a polypeptide. These included not only the site of specific proteolytic cleavages, but also amino acid trimming, amidation, glycosylation, and phosphorylation. The quest then evolved into the paradigm that within the secretory pathway, specific proteolytic machinery that results in the limited initial cleavage of proprotein and prohormone precursors must exist. Such machinery is dependent on the substrates, whose cleavages can occur within different organellear compartments, including the trans-Golgi network, cell surface, endosomes, and secretory granules. The sorting of proteins before processing enables a refined filtering process that allows for

controlled proteolysis in a precise spatiotemporal manner. This chapter considers insights from the field of Intramolecular chaperones and protein folding, and how it has provided a better understanding of the field of protein precursor processing, evolution of protease function, and biological regulation.

Identification of Proprotein Convertases

The identification of the yeast convertase kexin was a major stepping-stone in the prohormone field in the 1980s, as it provided the necessary molecular and genetic information to identify the first mammalian homologue, the prototypical proprotein convertase (PC), furin. The realization that the catalytic domains of mammalian furin and yeast kexin were homologs of a simple, broadly specific bacterial protease, subtilisin, was initially disappointing for the lack of a more exotic protease prototype, however it was quickly realized that the experimental tractability of subtilisin could be exploited to garner a great deal of insight into the biology of the subtilase superfamily. Almost every property of subtilisin, including its folding, catalysis, substrate specificity, pH profile, stability to oxidation, thermal and alkaline inactivation, has been altered via protein engineering. This plethora of information opened the floodgates to our understanding of the PC family.

Proprotein Convertases are members of the subtilase superfamily

The PCs are a subgroup within the subtilase superfamily, which covers all major domains of life, and represents the largest family of serine proteases after the chymotrypsin family [2].The subtilase family is defined by two features: i) a shared common fold, the subtilase fold, comprised of a seven-stranded β-sheet sandwiched between two layers of helices, and ii) its catalytic triad of an aspartate, histidine, and serine residue [3] (Figure 1.1 A). According to the Merops database, the subtilases constitute Family S8, and are divided into two subfamilies; the prokaryotic subtilisins are the archetype for subfamily S8A, and the yeast kexin the archetype for subfamily S8B [2]. The prokaryotic subtilisins are extremely stable and broadly specific, and thus are utilized extensively in detergents, cosmetics, food processing, and for research purposes ranging from synthetic organic chemistry to protein engineering [4]. Their widespread use and importance generated significant momentum to gather extensive biophysical, biochemical, and structural information, and has made the prokaryotic subtilisins the prototype of the subtilase superfamily. Among the prokaryotic subtilisins, subtilisin E (*B. subtilis*), subtilisin BPN (*B. amyloliquefaciens*), and subtilisin Carlsberg (*B. licheniformis*) have been most extensively studied [5]. Prior to 2003, when the crystal structures of furin [6] and kexin [7] were published, crystallographic data of the prokaryotic subtilases served as the only structural template from which to extrapolate hypotheses about the eukaryotic PCs [8, 9]. The high-resolution X-ray structures of furin [6] and kexin [7] have transformed

our understanding of the basis of remarkable specificity displayed by eukaryotic PCs [10] when compared with their prokaryotic counterparts that display a remarkably broad specificity. Simultaneously, they may potentially provide us with the means to better understand the structural and functional evolution of subtilases within a cellular context. For example, carboxy-terminal to the protease domain, PCs have an additional P-domain that is essential for catalytic function. Experimental studies demonstrate that the P domain, which can fold independently, contributes to the folding and thermodynamic stability of the catalytic domain through reciprocal hydrophobic interactions. The X-ray structure shows the P-domain is a true polypeptide domain with a closed barrel of β-strands that makes no contributions to the active site or specificity pockets. The P-domain is then followed by isozyme-specific elements that are primarily involved in localization. Both of these domains are absent in the bacterial prototypes (Figure 1.1 A and B), with kumamolysin representing the only bacterial subtilisin that contains a C-terminal extension that is similar to PCs. Thus, the ability to now compare structural information between the prokaryotic and eukaryotic members of the subtilase family has been instrumental in furthering our understanding of the structural and functional evolution that underlies the remarkable specificity of the eukaryotic PCs in comparison to their promiscuous prokaryotic counterparts [10].

To date, nine secretory serine proteases that are members of the mammalian PC family have been identified. Seven of these – furin, proprotein convertase 1/3 (PC1/3), PC2, PC4, PC5/6, and PC7/8 and paired basic amino

acid cleaving enzyme 4 (PACE4) – constitute the kexin subfamily, and activate a broad array of cellular and pathogenic precursor proteins by cleaving bonds at paired basic amino acid residues[11, 12] (Figure 1.2). The additional two, subtilisin-kexin isozyme 1 (SKI-1) and proprotein convertase subtilisin kexin 9 (PCSK9) belong to the pyrolysin- and proteinase K-like clades of the subtilisin family, respectively, and are unique in both their preferred cleavage motif and proteolytic activity [13]. SKI-1 cleaves client proteins after non-basic residues and regulates cholesterol and lipid homeostasis[14, 15], while PCSK9, to our current knowledge does not cleave any substrates other than itself, and rather binds to the low-density lipoprotein receptor to induce its internalization and degradation[12, 16]. In all cases, residues within the substrate-binding pocket largely define the substrate specificity of subtilases. In bacterial enzymes, the active site is lined by large, hydrophobic groups [17], while that of kexin and the PCs contain a large cluster of negatively charged amino acids that facilitates recognition and cleavage at the C-terminus of multiple basic amino acids [6, 10]. The larger structural organization of the subtilases has likewise been conserved. All members contain a signal peptide, a propeptide, and a subtilisin-like catalytic domain; additional C-terminal modules, including the P-domain, have been acquired over evolutionary time, however the core modules have been conserved from the prokaryotic subtilisins to the yeast kexin to the metazoan PCs.

Concept of Intramolecular Chaperones

The importance of the homology between the bacterial subtilisins and eukaryotic PCs is highlighted when one considers the unique propeptides of these proteases; analysis of the cDNA for subtilisin E suggested that is was synthesized as a zymogen, with a 77 residue N-terminal propeptide between the signal sequence and protease domain [18]. Two observations are particularly illuminating with respect to the role of the propeptide. First, when the propeptide is deleted genetically, the resulting protein is expressed robustly, but only folds to a stable, yet catalytically inactive, molten globule intermediate, indicating a role for the propeptide in generating a properly folded, active protease. Secondly, when active subtilisin is denatured using chaotropes, the unfolded protein is unable to refold into a catalytically active native state upon removal of the chaotrope [19] unless the propeptide is added to the folding reaction [20], further implicating the propeptide as essential for correct folding. Similar requirements have since been described in a variety of proteins, including proteases from the serine, cysteine, aspartyl and metalloproteases [21, 22] and non-proteases from prokaryotes (e.g.: α-lytic protease [23-25]; elastase [26, 27]), eukaryotes (e.g.: carboxypeptidase Y [28-31]; proteinase A [32]; procathepsins [33-35]), archea, and viruses (e.g.: endosialidases [36, 37]), suggesting that propeptide-dependent folding mechanisms may have emerged via convergent evolution [21, 38]. To distinguish them from the more commonly known molecular chaperones [39-41], such as GroEL and the HSPs, Inouye and colleagues coined the term Intramolecular Chaperone (IMC) [38] to denote propeptides that are part of the

primary sequence of a protease, and are indispensable for the correct folding of their cognate catalytic domains. Unlike classic chaperones, IMCs are highly substrate specific, and are true single-turnover foldases [42]. Following folding, the IMC is proteolytically degraded, effectively destroying part of the folding information, and making the process irreversible. Furthermore, rather than simply providing an environment to prevent misfolding, IMCs actually accelerate the rate of folding [5]. This is necessary due to the fact that the native state is in a kinetically, rather than thermodynamically, stable state [24]. Subtilisin remains correctly folded because the rate of unfolding is extremely slow; as overall thermodynamic equilibrium favors the unfolded state, and the equilibrium constant is equal to the ratio of the rates of folding and unfolding, the rate of folding must be even slower. Therefore, the IMC acts as a folding catalyst, accelerating to the rate of folding ~10^6-fold [43]. Crystal structures of subtilisin, complemented by genetic and biochemical analyses, have helped establish the putative mechanism of IMC-mediated protein folding, which lowers the energy barrier between the extremely stable molten globular intermediate and the native state to allow productive folding by interacting with two surface helices of the protease and providing a nucleus for folding [24, 43-47]. Subsequent to folding, the propeptide is removed via two spatiotemporally distinct endoproteolytic cleavages, each of which have a different pH optimum [48, 49], resulting in the maturation of the zymogen into enzymatically active subtilisin [50, 51]. Between autoprocessing events, the propeptide also functions as an inhibitor of its

cognate catalytic domain, thus additionally functioning as a regulator of enzymatic activity [50], a role that will be discussed in detail in a later section.

Divergent sequences, same structure

As discussed above, the extensive prior work on subtilases and other proteases has firmly established that IMC-mediated folding is a kinetically driven process [52]. The high degree of conservation of function across unrelated protease families suggests that evolution has independently converged on a common solution on multiple occasions [21]. Given this fact, it is not surprising that even among homologous families, sequence conservation within IMC domains is significantly lower than that observed between cognate catalytic domains [38]. Nonetheless, analysis of the sequences of subtilisin IMCs have highlighted several key characteristics, which after subsequent experimental testing, have been shown to be critical for function. First, two hydrophobic motifs, N1 and N2, are conserved across subtilases [53], and can be likewise identified within the IMCs of aqualysin and Pleurotus ostrearus proteinase A inhibitor 1 [54]. Interestingly, a computationally designed peptide chaperone, ProD, was forced to diverge from the IMC of subtilisin, yet sequences within N1 and N2 maintained a high degree of conservation [55]. Similarly, random mutagenesis of residues within the IMC, coupled with functional screening, demonstrated that while there was a high tolerance for mutation in general, mutations within the N1 and N2 motifs were often deleterious [56, 57]. Taken together, these observations suggest that these motifs are responsible for nucleating folding,

while the other, less well-conserved motifs may play a role in functional specificity.

The second piece of information we can glean from a comparison of the sequences of IMCs is that IMCs are more highly charged than their cognate catalytic domains [38]. While 12% of the amino acids in the protease domain of subtilisin E are charged, 36% of the residues of its IMC are charged, which directly complement a pocket around the substrate-binding site [47]. This asymmetric distribution in charge has likewise been observed in α-lytic protease (IMC 31%, catalytic 10%), carboxypeptidase Y (IMC 30%, catalytic 20%), and proteinase A (IMC 31%, mature 20%). Why the difference in charge between the IMC and catalytic domains has evolved remains unknown, but further research into this phenomenon promises to be insightful into the nature of kinetic barriers in folding pathways.

As may be inferred from the conservation of the three-dimensional fold of the IMCs of subtilisins in the absence of sequence conservation, protein structure evolves more slowly than protein sequence. While protease domains require maintenance of the precise spatial orientation of the catalytic triad, their cognate propeptides are under no such pressure; How these variable IMC domains can mediate the correct folding of structurally conserved catalytic domains is difficult to reconcile, but some insight is offered by the fact that despite significant variability in sequence, IMCs adopt similar structures. This suggests that evolution may favor structural conservation over sequence conservation in the case of these unique domains.

Folding to kinetic stability

The wealth of biochemical and structural information about the bacterial subtilisins amassed over the past decades have allowed for a detailed characterization of IMC-mediated folding. Eloquent and detailed descriptions of mature subtilisin E and subtilisin E in complex with its IMC have been published previously [46, 47, 58], as well as the structure of an active site mutant ($S_{221}C$) [46, 47, 59], all of which have provided snapshots of the transition from the unfolded state, to inhibition complex, to an active protease. Of particular interest to the current subject of review is how the IMCs of the subtilisins facilitate the folding of their catalytic domains. As noted above, attempts to fold subtilisin without its IMC resulted in the formation of a stable, yet non-functional protein. This intermediate had a hydrodynamic volume between that of the fully folded and fully unfolded protease, and circular dichroism spectra of the intermediate indicated a well-defined secondary structure, but a lack of tertiary packing. Similar results were seen in studies with α-lytic protease [60]. These studies suggested that the protease was able to fold to a kinetically-trapped, molten-globule intermediate [61] in the absence of its IMC, but a high energy barrier between the molten globular intermediate and transition state prevents folding to the native state. Addition of the IMC lowers this barrier to allow folding [24, 43-45]; therefore it was proposed that the IMC functions to overcome a kinetic barrier on the folding pathway.

To further dissect the role of the IMC in folding, refolding studies of full-length subtilisin E and subtilisin BPN were carried out using a catalytically

attenuated mutant ($S_{221}A$) [43, 62]. Equilibrium unfolding of the fully folded complex was monitored via changes in circular dicroism and fluorescence spectroscopy, and were noted to follow a three-state unfolding curve; most of the tertiary structure and part of the secondary structure unfolded in an initial cooperative step, while the remaining secondary structure followed a less cooperative second transition. Notably, this second unfolding transition was similar to the unfolding transition of the molten globule intermediate, and at higher concentrations of denaturant, the complex likewise displayed properties similar to the molten globular intermediate. Taken together, the three-state equilibrium unfolding transition suggested that the polypeptide first folds to a molten globule like state, then in the late stage of the folding pathway, transitions to the native state with the assistance of the IMC.

The final set of studies to define the role of the IMC in folding was to establish the kinetic parameters of IMC-mediated folding. Using the method of rapid dilution for fast refolding, Eder and Fersht demonstrated that the folding of subtilisin BPN follows two-state kinetics, with the IMC accelerating the kinetics of folding at least five orders of magnitude [43]. Furthermore, while the native state is approximately 1 kcal/mol less stable than the unfolded state [24, 62], the difference in energy of the transition state in the presence and absence of the IMC was estimated to be > 7.5 kcal/mol [21]. Therefore, folding with the assistance of an IMC appears to be a mechanism by which the protease can fold to a kinetically trapped, thermodynamically unstable native state.

The question of why a protein family would evolve to fold to a kinetically stable state versus a thermodynamic one is an interesting one, and provides some insight into the larger so-called “protein folding problem”. Conventional wisdom suggests that the primary sequence of a protein contains all of the information necessary to direct folding through a unique energy landscape[63, 64]; the challenge has been to understand how the amino acid sequence defines the folding pathway, and how this has been shaped by evolutionary forces. One approach to addressing this challenge is a comparison of homologous proteins that differ in one key feature; an analysis of all available sequences of bacterial serine proteases reveals two distinct subfamilies, the intracellular serine proteases (ISPs) and extracellular serine proteases (ESPs) [65]. ISPs and ESPs share a high level of sequence, structure, and functional homology, however ISPs lack the classical subtilisin propeptide. In fact, in contrast to ESPs, which absolutely require their propeptide to fold to a kinetically stable state, ISPs fold to a thermodynamically stable state independent of a domain homologous to the propeptide domain [66]. Furthermore, while the secondary structures of ISPs and ESPs are very similar and have similar thermostability, the pathways by which they fold to these states are different. Evolutionary analysis suggests the observed differences in folding pathways are mediated by differential selection of residues mapping to the surface of the protease and the interface formed between the protease and propeptide of ESPs. These results were the first to suggest that closely related subtilases can fold through distinct mechanisms and pathways, and that sequence can dictate the choice between propeptide-

dependent kinetic stability and propeptide-independent thermodynamic stability [65]. This is in contrast to the canonical theory that homologous proteins fold through similar pathways and folding transition states [67-69]. In considering the environment in which these proteases function, we can speculate as to the differing evolutionary pressures that may have driven this divergence. Subtilisin E is an ESP, secreted from the cell into a harsh, protease-rich extracellular environment to facilitate nutrient scavenging by the bacteria. As such, the presence of the propeptide domain allows it to fold to a kinetically trapped native state that is resistant to unfolding and proteolysis. In contrast, ISPs fold to a thermodynamically stable state very rapidly and without the requirement for a propeptide, and are responsible for a wide variety of intracellular processing events, regulating cellular functions such as DNA packing, genetic competence, and protein secretion [70-73]. As these processes require temporal regulation, both in production to initiate a pathway, and turnover to terminate it, the longer time of folding and kinetic stability conferred by the propeptide-dependent folding pathway of the ESPs may impede proper cellular function. This suggests that biological requirements, not just structural ones, may dictate folding pathways. While thermodynamically stable protein conformations are the norm, kinetically stable conformations may be selected for specific functional advantages [52] (Figure 1.3).

Interestingly, although the PCs more closely align with the ISPs in terms of biological function, their sequence and dependence on their propeptide for folding aligns them more closely with ESPs. We hypothesize that the choice of a

propeptide-dependent pathway was driven by the functional advantage of a scaffold upon which to build modulation of folding, interactions with transport machinery, and regulation of compartment-specific activation.

Regulation of activity

In general, cells utilize two mechanisms to control proteolytic activity; the first involves co-sorting of specific endogenous inhibitors into compartments where proteases are inactive, and spatially distinct from those containing active protease. The second requires proteases being synthesized as inactive precursors, which then become activated by a defined intra- or inter-molecular proteolysis. Intramolecular chaperone-mediated inhibition is a unique combination of these two strategies; IMC-dependent proteases are produced with their inhibitor as part of their primary sequence, which are then cleaved, but remains associated until the IMC: protease complex reaches the appropriate cellular location. As discussed above, subsequent to chaperoning folding, the IMC of subtilisin undergoes two autocatalytic cleavages; the nature of the interaction between the IMC and its cognate catalytic domain in between autoprocessing events was investigated using wild type and mutant IMC domains [48, 56, 74]; these studies revealed that the IMC functions as a slow binding competitive inhibitor of subtilisin. Generally, slow binding inhibition is characteristic of complexes where an initially weakly association is enhanced by a conformational change. In the case of subtilisin, the isolated IMC is intrinsically unstructured in its isolated state, but takes on a well-defined conformation upon binding to its cognate catalytic domain [42]. Similar studies in related protease

systems establish that the slow binding inhibition of IMCs is a characteristic theme. Since incorrect activation of a protease intracellularly can have devastating consequences [75], these inhibitory propeptides were hypothesized to have evolved to regulate the spatiotemporal activation of the protease.

Putting it all together: IMC regulated activation of subtilisin

Based on the information discussed above, a mechanistic model for the activation of subtilisin has been proposed (Figure 1.4). Prosubtilisin is secreted from the cell, and progresses through three distinct steps: folding, autoprocessing, and degradation of the IMC [49, 76]. Folding of prosubtilisin is a rapid process that requires the presence of its IMC, and proceeds through a partially structured folding intermediate [77]. The intermediate then undergoes a conformational change to produce the native-like folded catalytic domain that allows for autoprocessing [76]. The autoprocessed IMC-subtilisin complex is extremely stable, thus release and subsequent degradation of the IMC is the rate determining step of the maturation pathway, however once a single molecule of subtilisin is activated, it can facilitate further degradation of the IMC in trans, thus allowing for rapid, exponential activation of all available protease [66].

The final, rate determining step of subtilisin activation was demonstrated to be stochastic *in vitro* [66], consistent with the energetically unfavorable release of the tightly associated IMC from the active site. Further analysis of this stochastic behavior was achieved by varying solvent conditions to mimic the extracellular milieu into which the protease would be secreted, which perturbed

the release-rebinding equilibrium of the IMC, and in turn altered both the time and randomness of protease activation, suggesting that there is a carefully tuned balance between IMC structure, release, and protease activation. As regulation of proteolysis requires deterministic triggers, not stochastic chance, to ensure homeostasis, the selection of an IMC-dependent pathway likely represents a mechanism for regulation of protease activation.

Emergence of specificity from promiscuity

The Proprotein Convertases are the eukaryotic homologues of the bacterial subtilisins, and as such, while the level of complexity of their structure, function, and regulation has increased, they maintain distinct similarities. While the PCs have accumulated additional functional modules over evolutionary time, the core subtilisin structure, as well as behavior of the modules, has been conserved (Figure 1.2).

The catalytic domain

The subtilase superfamily is defined by its conserved catalytic triad and fold, the subtilase fold [3]. Crystal structures have revealed that while the catalytic domains of furin [6, 10] and yeast kexin [7] are highly conserved from bacterial subtilisins, there are important differences in the architecture of their individual substrate binding clefts. Structure and sequence comparisons suggests that, unlike bacterial subtilisins, the substrate binding sites of PCs contain a large cluster of negatively charged residues that facilitates the

recognition and cleavage at the C-terminus of multiple basic amino acids. However, the distribution and number of acidic residues in PCs differs between individual homologues, providing overlapping but not identical substrate specificity. Interestingly, PCSK9 shows unexpected, and even surprising, evolutionary and structural relationships with members of the PC family. The catalytic domain of PCSK9, although similar to furin and subtilisin, can only cleave its propeptide to form a non-covalent, proteolytically inactive complex that has a unique biological function [78]. Recently, profurin was shown to moonlight as a non-proteolytic chaperone by facilitating secretion of matrix metalloproteinase-28 through direct interactions with its propeptide[79]. Hence, evolutionary pressures appear to be the likely driving force behind the increased and evolving specificity of eukaryotic PCs compared with their promiscuous bacterial counterparts.

The P-domain and C-terminus

In addition to their well-conserved IMC and catalytic domains, the PCs have a unique additional domain, the P-domain [80], and a variable cysteine-rich C-terminus [81] that together are important for their catalytic activity, trafficking and localization. The specific function of the P-domains of the PCs have not been extensively investigated, however it has been hypothesized that it is involved in stabilization of calcium binding, and experiments where the P-domain has been mutated or truncated indicate it is indispensable for catalytic activity of the protease [82, 83]. It is noteworthy that kumamolisin, which is also an acid-activated protease, is the only bacterial subtilase that displays a C-domain similar

to the P-domain, although the precise function of this domain remains under investigation [84]. Our understanding of the role of the C-terminus of the PCs is only marginally more advanced, as it has been demonstrated that an α helical domain within the C-terminal tail of PC1/3, PC2 and PC5/6A is responsible for their targeting to the dense core secretory granules [85, 86].

Intramolecular Chaperones of Proprotein Convertases

Analogous to prokaryotic subtilisins, PCs encode a propeptide in their primary sequence that is required to chaperone folding of their catalytic domains [87-89], and is involved in the correct spatiotemporal regulation of their activity within the cell. Consistent with earlier observations across prokaryotic homologues, while the sequence conservation between IMC domains of bacterial subtilisins and PCs is low, the domains nonetheless adopt similar folds in complex with their cognate catalytic domains. Interestingly, based on the solution structure of the isolated mouse PC1/3 IMC [90, 91], the IMCs of PCs appear to be at least partially structured in the absence of their catalytic domain, unlike the bacterial subtilisins [42]. The structure reveals that the IMC forms a well-ordered core, consisting of two α-helices packed against a four-stranded antiparallel β-sheet. This interface is stabilized by a complementary surface created through a series of conserved hydrophobic residues [91]. Additionally, it was noted that that the C-terminus of the propeptide was largely unstructured in the isolated form, but became more structured upon binding to the catalytic domain, and thus may underlie the slow-binding behavior of the IMC. To further investigate the sequence-structure relationship across subtilases, a multiple sequence alignment

of IMCs from bacterial subtilisins and eukaryotic PCs was assembled; not surprisingly, the alignment revealed a high level of conservation in the N1 and N2 motifs, but little homology otherwise [92] (Figure 1.5).

Activation of the PCs

The activation pathway of PCs follows the same general scheme as that of bacterial subtilisins, with the propeptide first guiding folding, then undergoing a two-step autoproteolytic removal of the domain. This process has been most thoroughly described in furin [88]. Following translocation to endoplasmic reticulum by the signal peptide, the propeptide guides folding of its cognate catalytic domain. The propeptide is then rapidly ($t_{½}$ ~ 10 min) excised by cleavage at the C-terminus at the consensus cleavage site (–Arg-Thr-Lys-Arg_{107}), yet remains non-covalently associated in the heterodimeric propeptide: protease complex [93]. Blocking this cleavage by mutation at the active site results in an accumulation of the folded, unprocessed intermediate in the ER/Golgi intermediate compartment (ERGIC), suggesting that specific components of the trafficking machinery are able to detect correct formation of the IMC: protease inhibition complex to regulate their exit from this compartment. Following propeptide excision, the complex exits the ER, with the propeptide still bound as a potent inhibitor, and sorts to the TGN, where a second, slow ($t_{½}$ ~90 min) autocatalytic cleavage at a second, internal site within the propeptide (-Arg-Gly-Val-Thr-Lys-Arg_{75}-) occurs to release inhibition of the catalytic domain. Like the bacterial subtilisins, the second cleavage is rate-limiting [49, 66], and is followed by the rapid dissociation of the cleaved IMC from the active site [94].

Interestingly, while all PCs, with the exception of PC2, undergo the initial autoprocessing step in the ER, the second, activating cleavage happens only when the complex has trafficked to the appropriate cellular location. Furin and PC7 undergo this second cleavage in the trans-Golgi network, PC5/6 and PACE4 in the TGN an/or cell surface, while PC1/3 and PC2 are only fully active in the secretory granules.

Environmental sensing via Intramolecular Chaperones

The observation that the folding and trafficking of the proprotein convertases was driven by the ordered, compartment-specific cleavages of their propeptides argued for the coevolution of these sequences with the emergence of the secretory pathway in order to tune their specificity via regulation of activation. The necessity of furin to reach the TGN before becoming active, coupled with an understanding of the unique environments of each compartment of the secretory pathway supported the hypothesis that there must be a compartment-specific signal, as well as an encoded sensor that recognizes and responds to that environmental cue. The two factors that vary across the secretory pathway are the pH and calcium concentration of each compartment. While calcium concentration is critical for activity of the PCs, it was not sufficient to trigger activation [89]. Instead, calcium functions to stabilize the catalytic domain and participates in specific interactions with the substrate [10]. However, decreasing pH was necessary and sufficient to activate furin; Anderson and

colleagues demonstrated that a pH of 6.0, corresponding to the pH of the trans-Golgi network, triggered the second cleavage of the IMC of furin, coincident with appearance of furin activity within this compartment [88, 89, 94]. It is noteworthy that the optimum pH for the primary and secondary processing in bacterial subtilisins are also different [76].

It is the role of the IMC as a potent inhibitor that has become of special interest with respect to the proprotein convertases. Bacterial proteases, such as subtilisin, are generally secreted, whereas eukaryotic proteases predominate intracellularly and are found in subcellular compartments. As a result, while prokaryotic IMC-dependent proteases simply need to recognize the difference between 'inside' and 'outside', the problem posed by the secretory pathway of the eukaryotic cell is much more complex. The compartments of the secretory pathway serve to segregate specific biosynthetic and catalytic functions within membrane-limited organelles. Such compartmentalization likely evolved from the necessity to optimize performance of individual metabolic pathways by providing unique environmental conditions. Accordingly, the PCs exert their biological activities in a tissue- and compartment- specific manner. This is reflected in the primary sequence and structure of eukaryotic propeptides. All propeptides of the subtilases share a common fold, their sequences have almost no recognizable similarity, thus despite their similarities in enzymatic activity, the evolutionary divergence of their propeptides may explain the diversity of biological roles of the proprotein convertases [92, 95]. While the cleavage sites of the subtilisin propeptide are nonspecific, the sequence of the primary C-terminal cleavage as

well as the organelle-specific secondary internal site contain the consensus basic motif and are identical to the preferred sites within substrates. Given the promiscuous specificity of bacterial subtilases as compared to the stringency of the PCs, the differences in sequence reflect the divergence of eukaryotic propeptides from their prokaryotic progenitors due to differences in cellular environment.

As discussed earlier, while the PCs are more similar to the ISPs in terms of cellular function, structurally, they more closely resemble the ESPs. While both ESPs and ISPs are produced as zymogens, requiring removal of the propeptide for activation, only the propeptide of the ESPs acts as an IMC, a feature that has been conserved within the PCs. Undoubtedly, the requirement for diversifying function of individual proteins imposes differential constraints on the evolution of sequence and structure, however an analysis of patterns of divergence suggests that in many cases, variations have been made to a common core of shared structures. In the case of PCs, we hypothesize that the choice of an IMC-dependent pathway provided the framework necessary to integrate folding, trafficking, and regulation of compartment specific activation.

As most enzymes are exquisitely pH sensitive, the pH of each secretory and endocytic compartment critically determines and regulates the coordinated biochemical reactions that take place within a given organelle [96]. Segregation of specific biosynthetic and catalytic functions within membrane-limited organelles likely evolved to facilitate the optimization of individual metabolic pathways by providing unique environmental conditions and allowing energy

storage in the form of electrochemical gradients [97]. Using changes in pH for regulation and signaling has several advantages on the level of spatiotemporal response, a fact that is made readily apparent when cellular pH gradients are disrupted. Protons are single subatomic particles able to diffuse rapidly and induce reversible chemical changes on a near-instantaneous basis; protonation of specific residues within a protein, then, serves as a way to link changes in proton concentration to a proteins changing local environment to modulate structural changes, unmask active sites, modulate interactions between interaction partners, or cause more global rearrangements. There are numerous examples of proteins using protonation as a signal, including ligand release from HLA-DR [98], a mediator the Bohr effect in hemoglobin [99], and interaction of the translocation domain of diphtheria toxin with the host cell membrane [100].

Understanding the underlying determinants of pH sensing has powerful implications for the understanding of the fundamentals of protein evolution, folding and regulation. As discussed above, studies on furin have demonstrated that its 83-residue propeptide not only helps to fold the catalytic domain in the ER, but also requires two pH-dependent and compartment specific cleavages for furin to traffic correctly and produce active protease [88, 89, 94]. Similarly, sorting and proteolytic maturation of prohormones in secretory granules is strictly pH-dependent [101]. The overall two-step activation model is common to both bacterial and eukaryotic subtilisins, but the discovery that there were differential pH requirements of each cleavage in the IMC domains of furin suggested

evolutionary pressure favored facile spatiotemporal regulation of PC activity by exploiting the pH gradient of the secretory pathway.

Based on the aforementioned observations and an understanding of the general mechanism of activation in the bacterial subtilisins, it was hypothesized that protonation of an intrinsic pH sensor was the driving force behind activation of the PCs. In many ways, the role of pH-sensors is analogous to those of calcium sensors, such as calmodulin, that have binding affinities tuned to detect transient increases in calcium concentrations elicited by extracellular chemical signals [102]. While several amino acids are able to be protonated, one stands out as a prime candidate for pH sensing at physiologic pH.

While the pKas of most charged amino acid side chains are either strongly acidic or strongly basic, the imidazole ring of the histidine side chain can act as a general acid in its protonated state, or a general base in its deprotonated state, with a pKa of ~6.0. Large-scale analysis of the PFAM database identified 6533 unique subtilases with annotated propeptides, with members from the subtilisin, kexin, proteinase K, pyrolysin and sedolysin clades; comparison of the histidine content of the propeptide and catalytic domain revealed that the propeptides of eukaryotes contained a higher percentage of histidine residues than their cognate catalytic domains, while no such difference was observed in prokaryotic homologues [95] (Figure 1.6). A similar analysis of the domain content of other amino acids did not show similar trends, suggesting that the histidine enrichment is biologically significant. Interesting, an analysis of the outliers lends further support to the hypothesis that histidines may have been evolutionarily selected

for in order to recognize and respond to pH. Proteinase K and SKI-1, both eukaryotic proteases without enrichment of histidines in their propeptides, are both activated at neutral pH, thus relaxing the necessity for histidines. Similarly, kumamolysin and xanthomonalisin are members of the sedolysin family, which are found in acidic environments; as the intracellular pH is near neutral, the ability to sense the acidic extracellular environment has likely driven histidine enrichment to allow pH sensing. To ensure that these observations were not unique to subtilases, and were in fact related to the need for pH sensing, the cathepsins, a family of pH-activated cysteine proteases, and caspases, a family of cytosolic proteases, were also considered. As expected, while a bias for histidines was found in the propeptides of cathepsins, no such bias was observed in the caspases, further supporting the hypothesis that the functional requirement for histidines has been evolutionarily driven by the need to sense pH to regulate activation [95].

The now well-established fact that the activity of the PCs is regulated by their N-terminal propeptides, coupled with the histidine enrichment in this domain, further implicated the propeptide in the pH-dependent activation of the PCs. In a series of elegant swapping experiments, it was demonstrated that all of the information required for pH sensing is encoded in the propeptides of the PCs, and pH-dependent activation of the protease domains can be reassigned in a propeptide dictated manner (Chapter 2, [92]). The isolated propeptides of furin and PC1/3 are both well structured at pH 7.0, but begin to lose secondary structure as pH is lowered, undergoing a cooperative, sigmoidal unfolding to a

less well-folded state at pH 5.0; notably, the midpoint of these transitions coincided with the pH of activation of their cognate catalytic domains. Furthermore, when the propeptides of furin and PC1/3 are swapped both *in vitro* and *in vivo*, the pH at which their cognate catalytic domains are able to be activated is reassigned. Based on this finding, it was hypothesized that the propeptides of PCs have evolved to allow the exploitation the pH-gradient of the secretory pathway via an enrichment of histidines as a mechanism to regulate their activation within specific organelles. Further support for this hypothesis was lent by the identification of a conserved pH-sensing histidine within the propeptide of furin [103]. This histidine is adjacent to the secondary cleavage site, nestled in a solvent accessible pocket lined by hydrophobic residues. While in the near-neutral environment of the ER, the histidine is deprotonated, and thus behaves as a hydrophobic residue, stabilizing the packing within this pocket, and keeping the pH-sensitive loop protected against cleavage. However, upon entry into the acidic environment of the TGN, the histidine is protonated; as a result, the imidazole side-chain becomes polar, disrupting the packing of the pocket, and driving a local unfolding that exposes the secondary cleavage site, allowing rapid degradation and release of the propeptide from its cognate catalytic domain (Chapter 3, [104]).

As noted, the pH sensing histidine identified within the propeptide of furin is absolutely conserved in all PCs, however other members of the family are activated in different organellear compartments, and thus, different pH (Figure 1.7). For example, PC1/3 undergoes primary autoprocessing in the ER, and

transits the secretory pathway to the TGN like furin, but only becomes fully active in the acidic environs of the mature secretory granules; therefore, there must be additional determinants of pH-dependent activation. While we are only just beginning to understand how the differing pH sensitivities of the individual PCs is encoded, closer inspection of the sequence and structure of the propeptide of PC1/3 yields interesting insight; the conserved pH sensor, which is solvent accessible in furin, is more buried in the propeptide of PC1/3, and is at least partially overlade by a second histidine that is implicated in pH sensing. This second histidine may as a “gatekeeper”, modulating the solvent accessibility of the primary pH sensor, and thus tuning the pH dependent activation of the protease, a point that has been illustrated both mechanistically (Chapter 4, [105]) and via the demonstration that the pKa of the pH sensor in the PC1/3 propeptide is acid shifted relative to that of the pH sensor of furin [106]. Comparison of the measured pKa values of the pH sensing histidines within the propeptides of furin and PC1/3 with the computational predictions based on the solution structure of PC1/3 suggest that the mechanism of pH-dependent activation demonstrated in furin is conserved in PC1/3, and modulation of the pKa of the pH-sensing histidine allows for differences in pH- and organelle-specific activation. Again considering secreted proteases more broadly, these observations suggest an importance for spatial juxtapositioning of titratable groups in specific regions of proteins in order to maintain tight spatiotemporal control of their activation, an evolutionary adaptation that was likely concomitant with the emergence of

compartmentalization and specialization allowed by the secretory pathway within eukaryotic cells.

Implications of IMC-mediated regulation of the Proprotein Convertases

Intracellular pH regulates myriad processes in the cell, and as a result perturbations to pH homeostasis often have severe consequences; lowering pH by 0.3-0.4 can trigger apoptosis [107], while an increase of 0.2-0.3 promotes cell proliferation [108]. Furthermore, dysregulated intracellular pH is hallmark of several disease states, including cancer and atherosclerosis [75, 109, 110]. As the PCs are exquisitely regulated by pH, and involved in processing of a wide range of substrates, alterations in cellular pH are likely to affect the function of these proteases. Perturbation of cellular pH through disease may cause activation of a PC prematurely, or prevent activation all together. For example, excess lactic acid production and ATP hydrolysis as a result of the hypoxic tumor microenvironment causes acidification, which may drive activation of furin in the ER or alter trafficking to the cell surface, further upsetting homeostasis via unregulated proteolysis [111].

If there has been an evolutionary tuning of pH dependent activation within the propeptides of the PCs, and secreted proteases as a whole, the issue of “mistuning” becomes an interesting possibility to consider. The most obvious situation would be the loss of the conserved pH sensor, yielding either an inactivatable or constitutively active protease, depending on the type of mutation;

however more interesting is the possibility of mutation of one of the surrounding residues. Alterations within the local environs may alter the protonatability of the pH sensor and/or additional “tuning” residues. Recalling the role of the propeptide as a foldase, an interesting implication of IMC-mediated folding of the subtilases is that steric information is imprinted upon the protease by the propeptide, a phenomenon called protein memory. Shinde and colleagues demonstrated that an identical sequence for subtilisin gave rise to two different three-dimensional structures, with different secondary structures, thermostability, and substrate specificity dependent on whether its folding had been chaperoned by the natural propeptide, or one that had been genetically altered by a single point mutation [112]. It therefore follows that mutations within the propeptide can also alter other critical features of the PCs, which have significant implications for proprotein processing and organismal homeostasis. This is exemplified by the results of several single nucleotides polymorphisms within the propeptide of PC1/3, all of which have been demonstrated to impact structure, stability, and substrate specificity [113] (Chapter 5, Williamson et al., in preparation) of the protease.

In conclusion, the importance of appropriate proteolytic processing of substrates within the secretory pathway cannot be understated, thus as we continue to understand the interplay between the proprotein convertases, proton concentration and propeptide encoded pH sensors, we hope to gain better insight into how things go wrong in the disease-state, and how we can better approach situations to restore homeostasis.

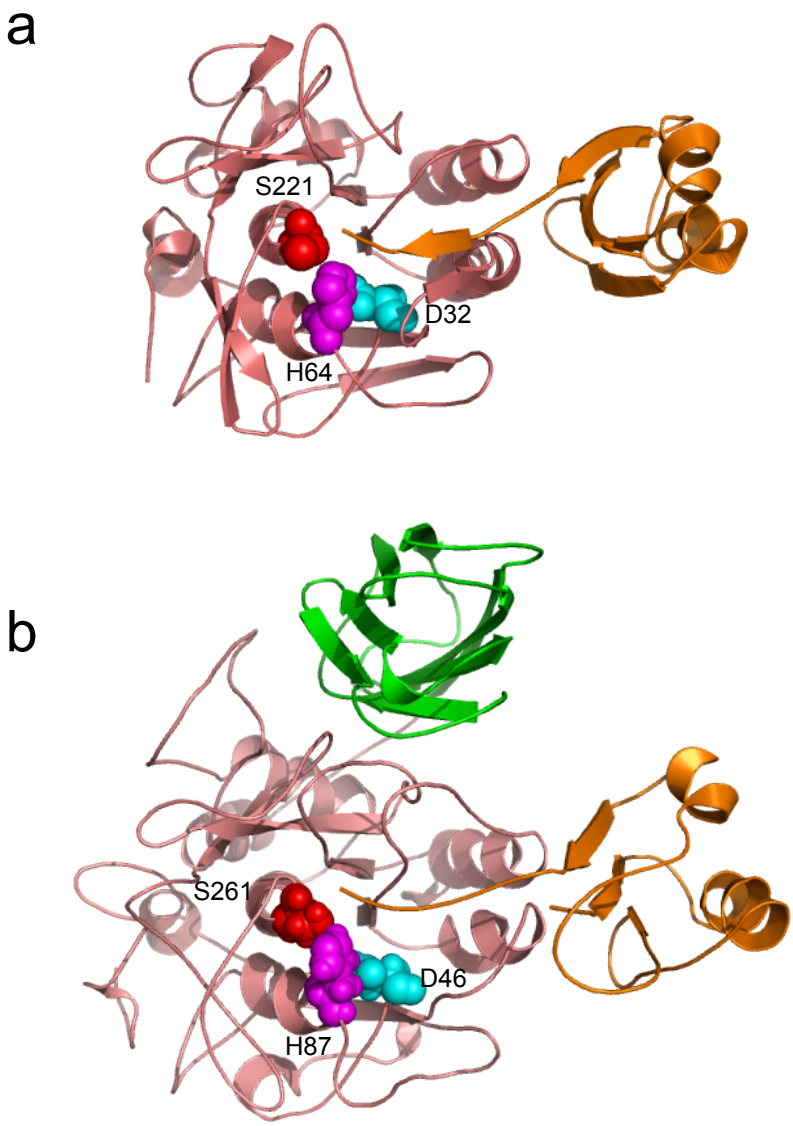


Figure 1.1: Structural conservation of prokaryotic and eukaryotic subtilases.
All subtilases have a well-conserved subtilisin-like catalytic domain (red ribbon), catalytic triad (depicted by space filling side chains, numbered with reference to beginning of catalytic domain), and propeptide (orange ribbon) **A)** Structure of subtilisin E (1SCJ) **B)** Structure of furin (1P8J) with its P-domain (green ribbon).

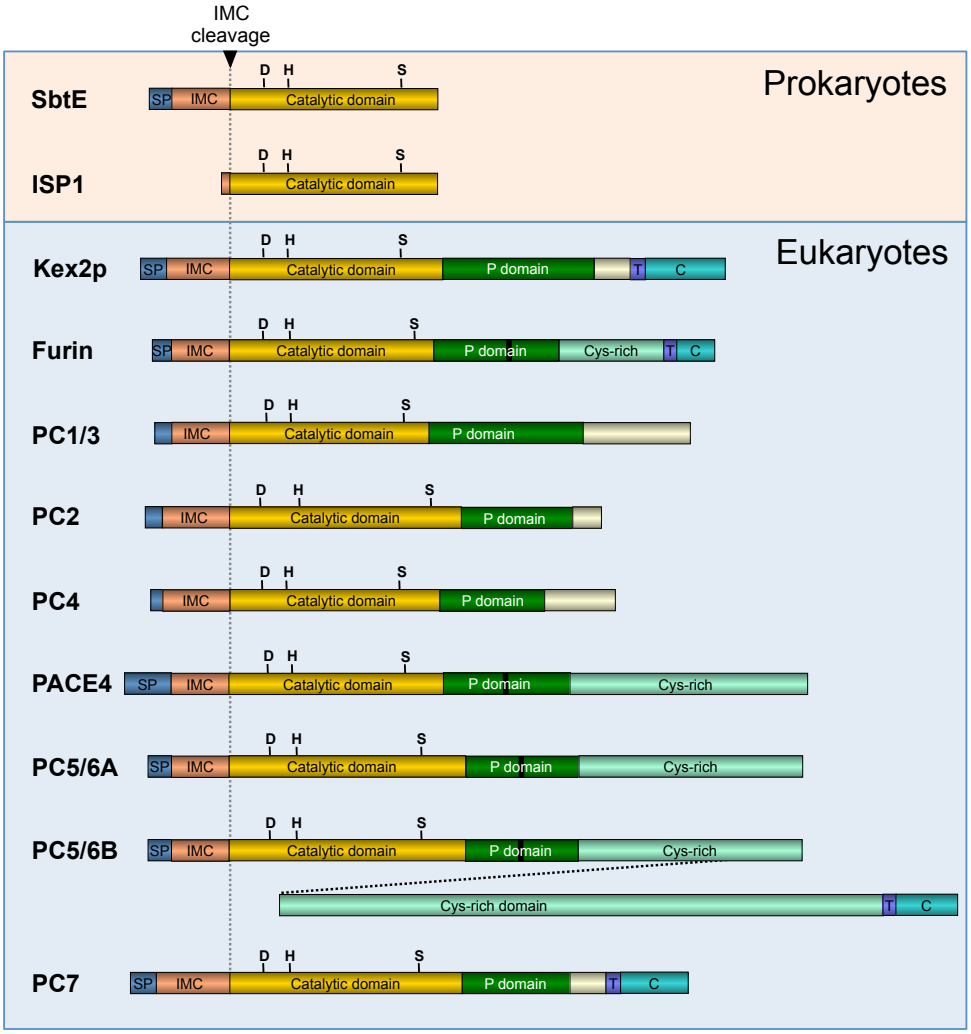


Figure 1.2:Domain organization of Subtilases.
Prokaryotic subtilases are in the orange box, eukaryotic subtilases in the blue box. The grey dashed line indicates the site of IMC cleavage to initiate activation. Residues of the catalytic triad are indicated. SP, signal peptide; T, transmembrane domain; C, cytoplasmic domain.

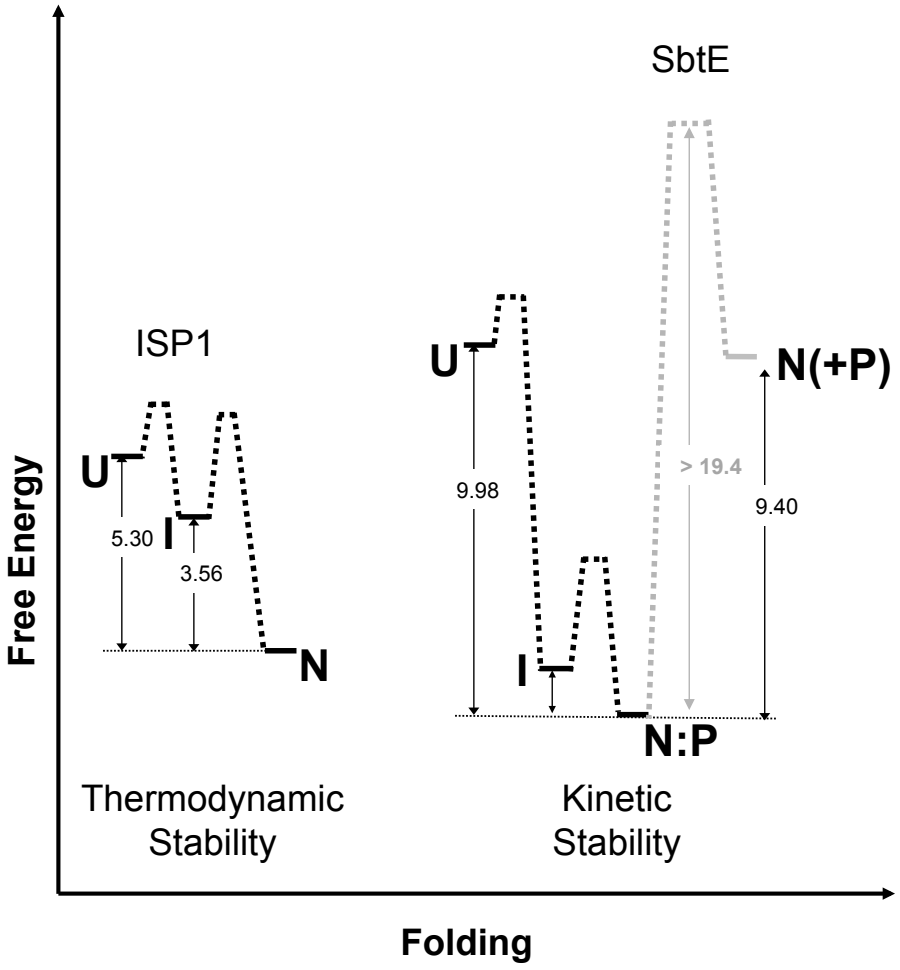


Figure 1.3: Free energy diagram of folding pathways of a representative ISP and ESP.
Left: Unfolded ISP1 (U) spontaneously folds to its thermodynamically stable native state (N) through a partially folded intermediate (I). The free energy difference between N and U is ~5.3 kcal/mol [65]. **Right:** Unfolded SbtE (U) undergoes rapid folding and autoprotolysis to give a thermodynamically stable IMC: SbtE complex (N: P) via an intermediate state (I) [114]. The activation energy for the spontaneous release of the IMC from this complex to release active protease (N + P) is energetically unfavorable, and is a stochastic process (Subbian 2005 JMB) [66]. The high activation energy barrier kinetically traps folded SbtE in its native state.

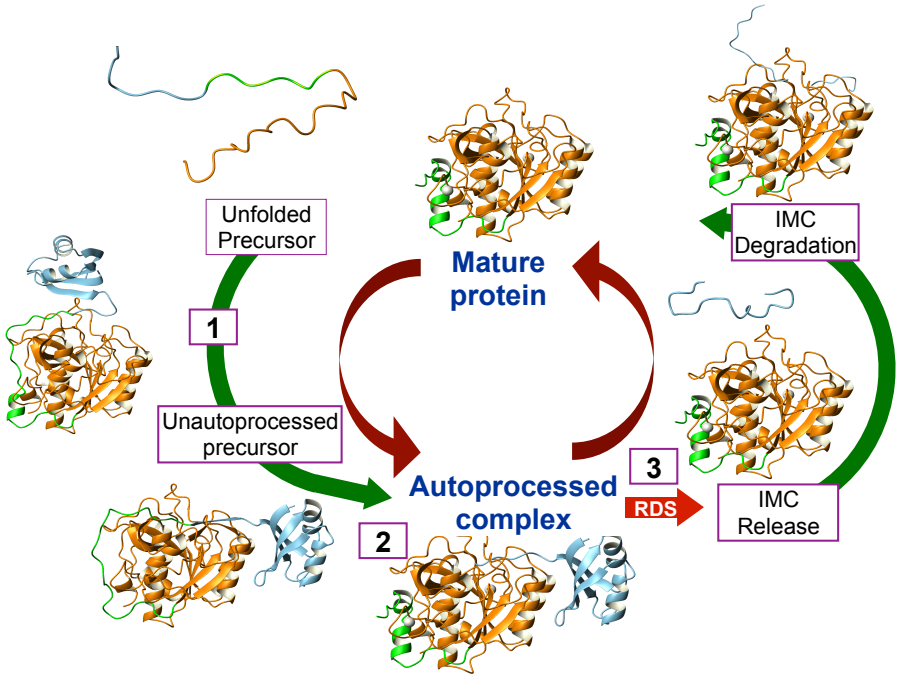


Figure 1.4: IMC-mediated pathway of subtilisin maturation.
The N-terminal helix of the protease is shown in green, the protease in orange, and IMC in blue. Maturation of subtilisin occurs in three stages: **1)** Folding of the unfolded precursor polypeptide chain to a structured, unautoprocessed precursor with the assistance of the IMC; **2)** Once the active site is formed, the precursor autoproteolyzes to an inhibited, autoprocessed complex; **3)** Release and degradation of the IMC from the autoprocessed complex generates active protease that can subsequently activate other autoprocessed complexes in *trans*. Activation is the rate-limiting step to maturation.

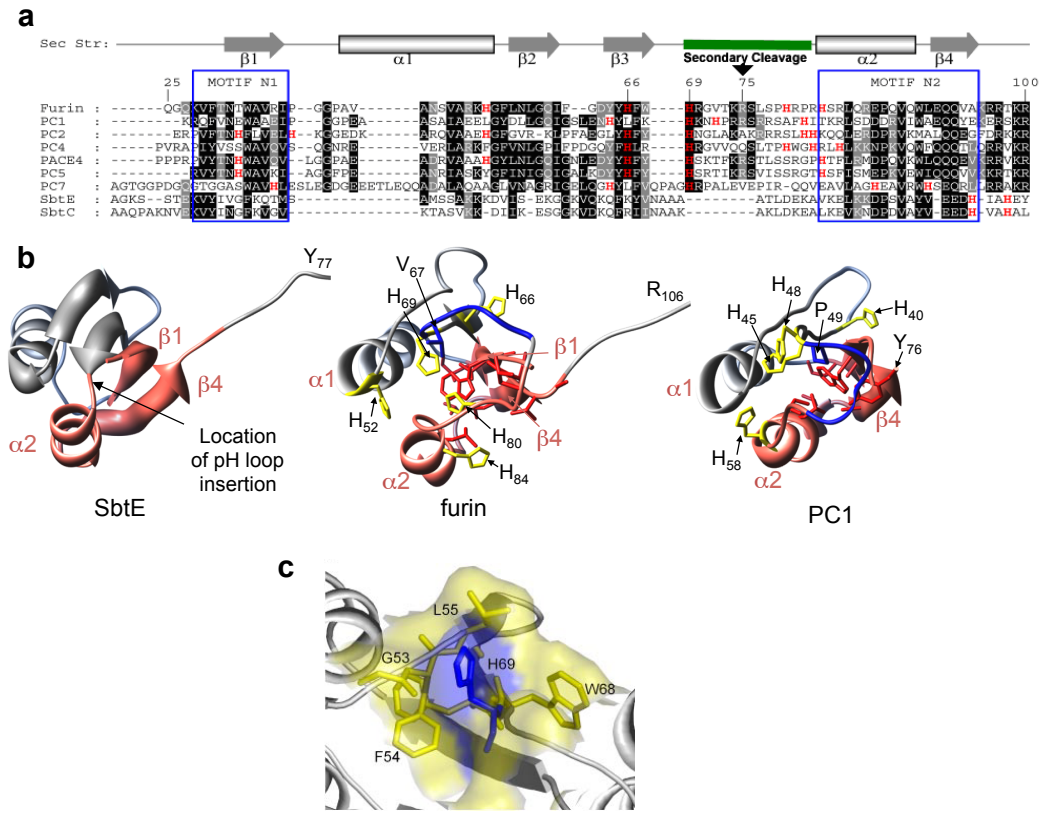


Figure 1.5: Comparison of IMCs from PCs and bacterial subtilisins.
A) Multiple sequence alignment of the sequences of the IMCs first seven PCs with IMCs of subtilisin E (SbtE) and subtilisin BPN (SbtC). Residues are numbered according to the furin sequence, which begins with Gln25. Histidines are highlighted in red. Predicted secondary structure and cleavage sites, indicated above the MSA, is based on the solution structures of the IMC of SbtE (1SCJ) [47] and PC1.3 (1KN6) [91]. Well-conserved motifs N1 and N2, the folding nucleation sites in subtilisin, are indicated in blue boxes. **B)** Comparison of the structures of IMCs of subtilisin E (left), furin (middle) and PC1/3 (right). The secondary cleavage site is indicated in blue. Key residues are indicated, with histidines highlighted in yellow. **C)** Surface representation of the solvent-accessible hydrophobic pocket formed by Gly53, Phe54, Leu55, Phe67 and Trp68 (blue), illustrating packing of His69 (red).

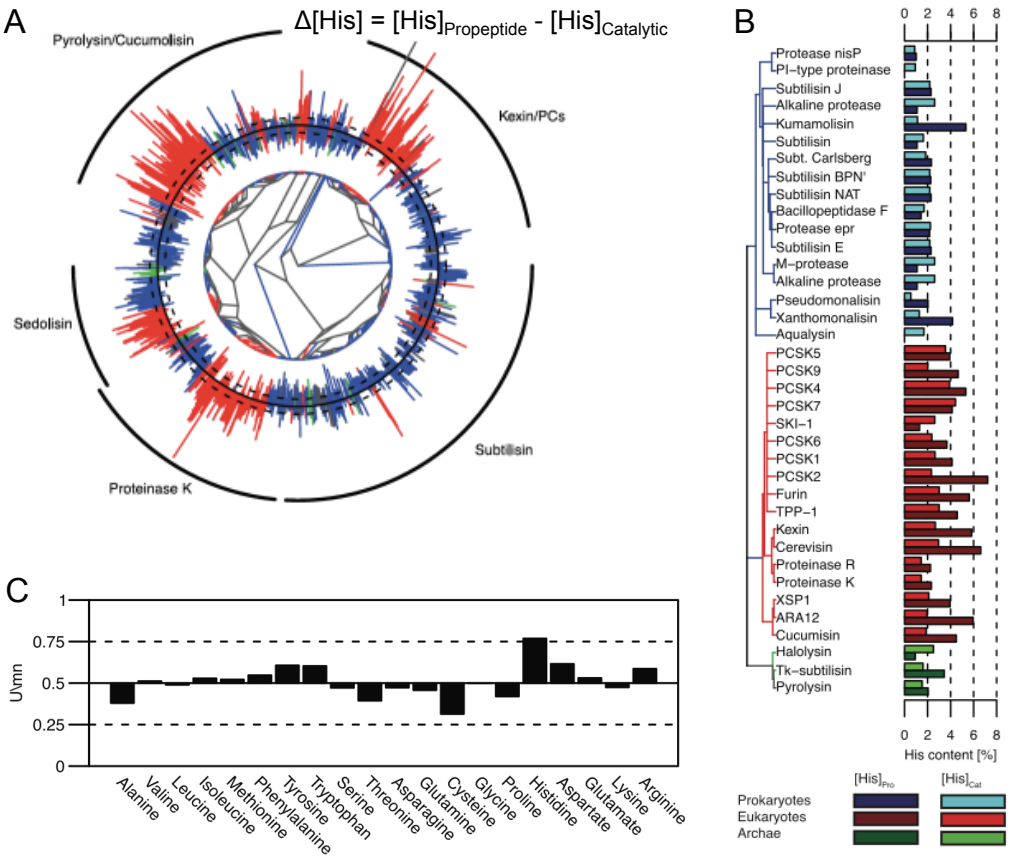


Figure 1.6: Histidines are enriched in propeptides of eukaryotic proteases. **A)** Circular phylogenetic tree of subtilases. Radiating lines indicate Δ[His] values, where Δ[His] = $[His]_{Propeptide}$ – $[His]_{Catalytic}$. The solid circle indicates Δ[His] = 0, and dashed circles indicate Δ[His] =+/- 1. Eukaryotic sequences are colored red, prokaryotic sequences blue, and archeal sequences green. Black arcs indicate clades of the subtilase superfamily. **B)** Bar graph showing $[His]_{Propeptide}$ and $[His]_{Catalytic}$ for selected subtilases. Prokaryotes are blue, eukaryotes are red, and archea are green. **C)** Effect size (U/mn) of differences between distributions of all 20 amino acids in prokaryotes versus eukaryotes. Values estimate the probability that a random sample of Δ[AA] in eukaryotes is larger than in prokaryotes. A value of 0.5 indicates equal distribution. Figure adapted from [95].

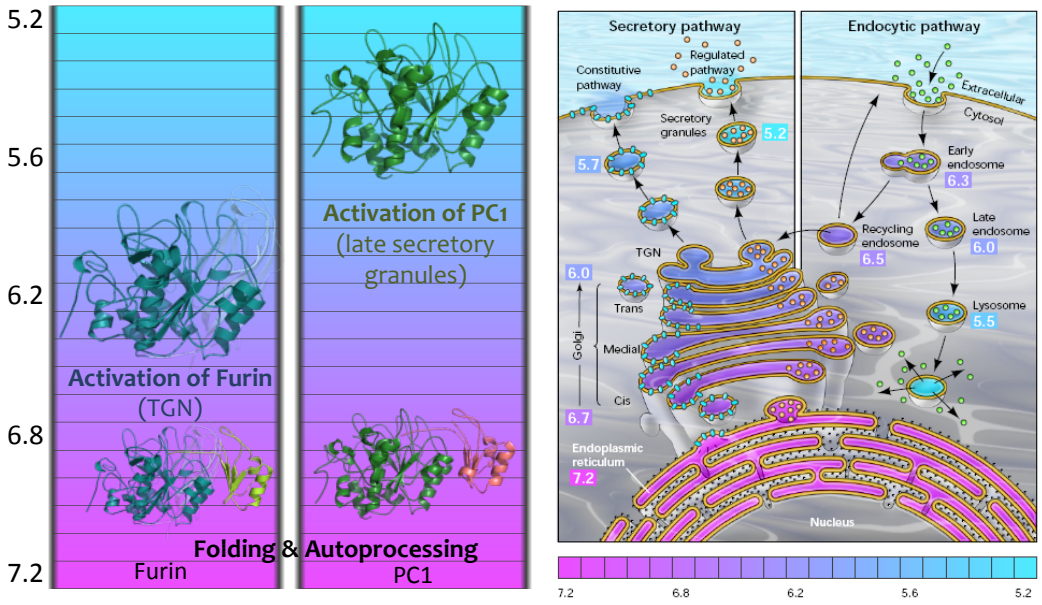


Figure 1.7: Furin and PC1/3 are activated at different pHs.
Left: Graphic representation of the approximate pH of the three processing steps leading to PC activation in furin and PC1/3. **Right:** Cartoon representation of the secretory pathway of a eukaryotic cell. The pH gradient of the secretory pathway is indicated by the transition from magenta (neutral pH) to aqua (cyan).

Chapter II: Propeptides are sufficient to regulate organelle-specific pH-dependent activation of furin and proprotein convertase 1/3

Stephanie L. Dillon[1,a], Danielle M. Williamson[1,a], Johannes Elferich[1], David Radler[1], Rajendra Joshi[2], Gary Thomas[3] and Ujwal Shinde[1]

[a]These authors contributed equally to this work

[1]Department of Biochemistry and Molecular Biology and [3]The Vollum Institute, Oregon Health and Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97229

and

[2]Scientific and Engineering Computing Group, CDAC, Pune University Campus, Ganeshkhind, Pune - 411 007 India

This chapter was originally published in The Journal of Molecular Biology in June 2012. Experiments were performed by SLD and DMW, JE helped with MD simulations, and DR with cloning. RJ provided the computational infrastructure. The manuscript was written by US and SLD, with contributions and editing provided by DMW, and GT.

Abstract

Proprotein convertases (PCs), furin and proprotein convertase 1/3 (PC1), cleave substrates at dibasic residues along the eukaryotic secretory/endocytic pathway. PCs are evolutionarily related to bacterial subtilisin and are synthesized as zymogens. They contain N-terminal propeptides (PRO) that function as dedicated catalysts that facilitate folding and regulate activation of cognate proteases through multiple-ordered cleavages. Previous studies identified a histidine residue (His^{69}) that functions as a pH sensor in the propeptide of furin (PRO^{FUR}), which regulates furin activation at pH~6.5 within the trans Golgi network. Although this residue is conserved in the PC1 propeptide (PRO^{PC1}), PC1 nonetheless activates at pH~5.5 within the dense core secretory granules. Here we analyze the mechanism by which PRO^{FUR} regulates furin activation and examine why PRO^{FUR} and PRO^{PC1} differ in their pH-dependent activation. Sequence analyses establish that while both PRO^{FUR} and PRO^{PC1} are enriched in histidines when compared with cognate catalytic-domains and prokaryotic orthologs, histidine content in PRO^{FUR} is ~two-fold greater than PRO^{PC1}, which may augment its pH sensitivity. Spectroscopy and molecular dynamics establish that histidine-protonation significantly unfolds PRO^{FUR} when compared to PRO^{PC1} to enhance autoproteolysis. We further demonstrate that PRO^{FUR} and PRO^{PC1} are sufficient to confer organelle sensing on folding and activation of their cognate proteases. Swapping propeptides between furin and PC1 transfers pH-dependent protease activation in a propeptide-dictated manner *in vitro* and in

cells. Since prokaryotes lack organelles and eukaryotic PCs evolved from propeptide-dependent, not propeptide-independent prokaryotic subtilases, our results suggest that histidine enrichment may have enabled propeptides to evolve to exploit pH-gradients to activate within specific organelles.

Introduction

Proprotein convertases (PCs) are endoproteases that mediate diverse regulatory and protective processes by controlled proteolysis of their substrates [88, 115, 116]. The PC-family includes seven mammalian Ca^{2+} dependent endoproteases: furin, PC1/PC3, PC2, PC4, PACE4, PC5/PC6, and PC7/LPC/PC8 [80, 88, 117-119]. More recently, SKI/S1P [14, 15] and NARC-1/PCSK9 [12] have also been identified as enzymes that share sequence similarity with PCs[13]. Structures of the catalytic domains of furin[6] and yeast kexin [7] have been solved using X-ray crystallography, and homology models for PCs that were derived from these structures provide the basis for substrate specificity [6, 7, 10, 120]. Although PCs potentially share overlapping cleavage specificity and function, each PC catalyzes limited proteolysis of proprotein and prohormone substrates at a pair of basic residues to excise bioactive proteins and peptides within specific compartments of the TGN/endosomal system that are characteristic of eukaryotic cells[13, 115, 121]. The necessity for such precise spatiotemporal cleavage of substrates mandates the activity of PCs to be likewise stringently controlled [122]. Dysregulation of PC activity has been observed in various diseases such as cancer [123, 124], obesity [125, 126], diabetes [127] and heart disease [128]. Consistent with these observations, small-molecule inhibitors of the constitutively expressed furin can inhibit cancer cell motility and invasiveness [129].

The activity of PCs is regulated by N-terminal propeptide-domains [89], which function initially as folding catalysts that facilitate folding of cognate

protease domains, and subsequently serve as temporary inhibitors that mask the protease active site [122]. Given their ability to chaperone single-turnover folding, these propeptides are often referred to as intramolecular chaperones (IMCs) [130]. IMCs constitute diverse, substrate-specific, single-turnover, energy-independent chaperones [25, 94, 103] whose primary sequences have diverged faster than their target client substrates [5] and are distinct from substrate-promiscuous, multi-turnover, energy-dependent inter-molecular chaperones [5, 122]. PCs are homologs of prokaryotic subtilases [3, 131, 132], proteins in which the roles of propeptides have been thoroughly investigated. Analysis of prokaryotic subtilases and other proteases suggests that propeptides evolved to regulate protease folding within harsh extracellular environments such as soil or vegetation [25, 62, 65, 66, 133, 134], and are absent from paralogues functioning in milder intracellular environments [65].

Subsequent to guiding protease-domain folding, propeptide-dependent subtilases undergo ordered proteolytic cleavages within their propeptide-domains. The first cleavage forms catalytically inactive propeptide: protease inhibition complexes wherein propeptides non-covalently bind to protease active-sites, while subsequent cleavages activate proteases by facilitating propeptide dissociation, enabling the now unmasked catalytic domain to cleave substrates *in trans* [13, 94, 103, 135]. While these obligatory cleavages are extracellular events in prokaryotes that delay onset of protease activity until after protein export, they control secretory pathway compartment-specific activation of substrate-specific eukaryotic proprotein convertases (PCs) [13]. Since eukaryotic

PCs evolved from propeptide-dependent, and not propeptide-independent prokaryotic subtilases [65], it is tempting to speculate that propeptides confer functional advantages through speciation, namely to regulate organelle-specific activation of secretory-pathway proteases, a complexity absent in unicellular prokaryotes that is essential to maintain physiological homeostasis within eukaryotic cells [97, 136, 137]. For example, the activation of furin is regulated in a pH-dependent manner as it transits the secretory pathway [94]. In the neutral pH in the ER, the propeptide is cleaved to form a stoichiometric propeptide: furin inhibition complex. Upon reaching the early Trans-Golgi network (TGN; pH 6.5) the furin propeptide (PRO^{FUR}) undergoes a second cleavage which removes the inhibitory propeptide and thus activates furin [94]. While PC1 transits the secretory pathway in much the same way, the PC1 propeptide (PRO^{PC1}) remains in a stoichiometric complex with the PC1 protease domain until undergoing its activating second cleavage upon reaching the dense core secretory granules (DCSGs; pH 5.5). A study by Feliciangeli et al.[103] demonstrated that in furin, mutating residue His^{69} in the propeptide to leucine blocks activation of the complex in the TGN while allowing for correct folding, while a $His^{69}Lys$ substitution results in accumulation of unprocessed furin precursor in the ER [103]. On this basis, they suggested that the His^{69} in the propeptide is not only important for folding of furin, but is also a vital pH-sensor that regulates furin activation in the pH of the TGN. However, mechanisms by which the His^{69} functions as a pH sensor in furin are unknown. Moreover, while the residue corresponding to His^{69} (in furin) is strictly conserved within the PC-family, PC1

and furin undergo their activating second cleavages at different pH within the TGN and DCSGs, respectively. This suggests that additional factors may play a role in regulating activation of the protease domains.

In this manuscript we demonstrate through various biophysical, biochemical, cell-based and computational approaches that the propeptide domains of furin and PC1 (PRO^{FUR} and PRO^{PC1}, respectively) contain sufficient information to confer organelle sensing on the folding and activation of cognate proteases. Circular dichroism spectroscopy as a function of pH establishes that the pH-dependent stability of propeptide domains coincides with the optimum pH for compartment specific activation. Monitored by ellipticity at 222 nm the PRO^{FUR} undergoes a transition in structure, the midpoint of which occurs at pH 6.5, while the midpoint in structural transition for PRO^{PC1} occurs at a lower pH (pH~5.5). Furthermore, swapping propeptides between eukaryotic paralogues —furin and PC1— transfers pH-dependent protease activation in a propeptide-dictated manner *in vitro* and in cells. Our results suggest that PRO^{FUR} and PRO^{PC1} encode information essential for regulating compartment specific activation of cognate proteases and that other residues in addition to the conserved pH sensor His^{69} are necessary to enable subtle differentiation in pH-dependent activation between furin and PC1. Using molecular dynamics simulations, we also demonstrate that histidine protonation leads to conformational changes in PRO^{FUR} but not in PRO^{PC1}. Together, our results provide insights into the structural mechanisms by which propeptides can regulate the pH-dependent activation of their cognate PCs.

Results

Eukaryotic propeptides harbor an internal cleavage site loop that is missing within their prokaryotic paralogues

To understand how eukaryotic propeptides can mediate compartment specific activation of their cognate protease domains, we compared sequences and structures of prokaryotic propeptides— subtilisin (PRO^{SUB}) and aqualysin I (PRO^{AQU}) —with eukaryotic propeptides— pro-protein convertase 1 (PRO^{PC1}) and furin (PRO^{FUR}). While several laboratories have analyzed the sequences and structures of propeptides, to date no detailed comparison between the sequences and structures of the propeptides of prokaryotic and eukaryotic proteins has been conducted. PRO^{AQU} was selected because unlike its intrinsically unfolded prokaryotic homologue PRO^{SUB}, PRO^{AQU} adopts a well-defined structure and chaperones folding of its cognate protease domain [138]. From the PC-family members, we selected PRO^{PC1} and PRO^{FUR} because despite significant sequence and structural similarity with prokaryotic orthologs (Fig.2.1A and 2.1B), they activate in different organelles along the proton-gradient of the secretory pathway, a complexity missing in prokaryotes. Furin is optimally active at pH 6.5, consistent with its role in cleaving proprotein substrates in the mildly acidic environment of the TGN/endosomal system. PC1 is optimally active at pH 5.5, consistent with its role in cleaving prohormone molecules in secretory granules.

Amino acids absent between residues 75-81 in PRO^{SUB} (red box; Fig. 2.1A) coincide with organelle-specific cleavage-sites [103] within eukaryotes (red loop; Fig. 1B). In prokaryotic subtilases, the secondary cleavage site is fairly promiscuous and presumably occurs in the flexible region between b_1 and a_1 (Fig 2.1B). Additionally, there are significant differences in residues 100-107 within the propeptide-domains between prokaryotic subtilisins and eukaryotic PCs. This C-terminal region harbors the primary cleavage site within propeptides and interacts with the substrate binding regions within cognate proteases to initiate activation. It is noteworthy that cellular substrates of PCs contain the consensus sequence [R/K]-X_n-[R/K]↓, identical to the primary cleavage site within propeptides [139]. Given the promiscuous specificity of bacterial subtilases when compared to the stringent substrate specificity of eukaryotic PCs, the differences between residues 100-107 reflects the requirement of PCs to cleave at highly conserved dibasic residues. This region reflects the divergence of propeptides from prokaryotes and eukaryotes to function with more cleavage specificity, likely due to the difference in cellular environment, namely, the inclusion of membrane bound organelles in eukaryotes [122].

PRO^{FUR} and PRO^{PC1} are rich in histidine residues when compared with PRO^{SUB} and PRO^{AQU}.

The pH within an organelle can dramatically affect the ionization states of charged residues in a protein sequence, by altering its structure, stability and function. Hence, we next analyzed the fold increase in amino acid residues within the propeptides and cognate proteases within prokaryotic and eukaryotic

subtilases using the averaged amino acid distribution within the Uniprot database as our baseline (Fig. 2.11C). The individual amino acid content for each family of propeptides and proteases were calculated and averaged. The contents of amino acids belonging to individual groups were added and divided by the sum of their content in the whole UniProt database (release 2011_12) to obtain the fold change as described in the Methods Section. Fold values greater than one (varying shades of red) indicate residue enrichment in propeptide domains within an individual group, values less than one (varying shades of green) indicate depletion of specific residues within propeptides, while a value of one (white) indicates no change. This graphical representation of the fold increase in specific groups of amino acid residues (Fig. 2.1C) demonstrates that the His content in PRO^{FUR} and PRO^{PC1} from eukaryotes is significantly greater that their cognate catalytic domains and prokaryotic paralogues. Furthermore, protease domains of prokaryotes are biased towards acidic and basic residues as demonstrated by Inouye and co-workers [5, 22, 38] which was hypothesized to enhance kinetic stability within their catalytic domains[65]. The average composition of proteins in the Uniprot database establishes histidine (2.27%) as the third least abundant residue, and is ~four-fold less than leucine (9.67%), the most abundant residue. While propeptide domains generally display a bias for charged and polar residues when compared to proteases[5], it is noteworthy that within subtilases, only eukaryotic propeptides are rich in histidine-content (Fig. 2.1C) when compared with prokaryotic propeptides and cognate catalytic-domains. Similar results are observed when propeptides within the PC-family are compared with

their cognate catalytic domains and prokaryotic orthologs (data not shown). Histidine is a unique residue because the pKa of its imidazole side-chain (pH~6.0) is close to physiological pH, and relatively small shifts along the proton gradient can change the net charge, and subsequently alter pH-dependent conformational stability of propeptides.

PRO^{FUR} regulates furin activation by acting as a pH-sensor that delays the internal propeptide cleavage until after the PRO^{FUR}:furin complex is trafficked to the mildly acidic TGN/endosomal system. Protonation of His^{69} forms a cleavable furin site at Arg^{75} which releases the bound propeptide from the catalytic domain [103]. Based on the demonstration of a histidine driven pH-sensor in PRO^{FUR} [103] and the histidine bias within PRO^{FUR} and PRO^{PC1} from eukaryotes when compared with PRO^{SUB} and PRO^{AQU} from prokaryotes, we propose that histidine protonation may regulate organelle-specific propeptide-release/degradation to regulate furin and PC1 activation in endosomal/lysosomal compartments elaborated in eukaryotic cells.

Circular dichroism spectroscopy demonstrates pH dependent structural changes in eukaryotic propeptides

Since the pKa (~6.0) of the imidazole side-chain of histidine is close to physiological pH, we next investigated whether small changes in proton concentration alter pH-dependent structural stability of propeptides in prokaryotes and eukaryotes. The secondary structures measured using circular dichroism spectroscopy measured at pH 7.0 demonstrates that PRO^{FUR} and PRO^{PC1} adopt structures similar to PRO^{AQU} and $PRO^{SUB\text{-}C}$ complexed to subtilisin

(Fig. 2.2A). Since isolated PRO^{SUB} is intrinsically unstructured[47], PRO^{SUB-C} structure was obtained by a difference spectra between the cleaved PRO^{SUB}:S_{221}C-subtilisin complex and mature subtilisin as described earlier[42].

The pH dependent structural stability of various propeptides was monitored by observing changes in negative ellipticity at 222 nm as a function pH (Fig 2.2B); as a representative example, we show the complete CD spectrum of PRO^{FUR} at the two ends of the pH range (pH 7.4 and pH 5.0) compared with a completely denatured PRO^{FUR} (Fig 1.2C). It is noteworthy that when the pH of the buffer is lowered from pH 7.4 to pH 5.0, PRO^{FUR} loses approximately 25% of its ellipticity at 222 nm when compared with the propeptide completely denatured in 8 M urea. Furthermore, changes in negative ellipticity at 222 nm as a function of pH (Fig 1.2B) suggest that the conformation of PRO^{FUR} tends to stabilize at approximately -2800 $deg.cm^2/dmol^{-1}$ under acidic conditions, but does not reach the ellipticity of completely unfolded PRO^{FUR} (approximately -20 $deg.cm^2/dmol^{-1}$). This suggests that the changes in pH do not result in complete unfolding and that PRO^{FUR} may adopt a partially folded molten-globule like state similar to that observed using NMR spectroscopy under acidic conditions [140]. The NMR data also suggest that PRO^{PC1} and PRO^{FUR} do not aggregate in their isolated forms.

When conformational changes of the propeptides as a function of pH are compared, it is evident that PRO^{PC1} and PRO^{FUR} unfold at different pHs, ~5.5 and ~6.5, respectively (Fig. 2.2B). Although the unfolding of PRO^{PC1} is not complete at pH 5.0, the structure of PRO^{PC1} at a pH below 5.0 was not analyzed because it is beyond the range of the buffering capacity of our system. While this prevents

the accurate determination of the mid-point of unfolding transition in the case of PRO^{PC1}, changing buffer systems to accommodate lower pH is problematic because diverse ions that can differentially influence structure, stability and/or activity of the propeptide and protease system. Nonetheless, comparing the folding transitions profiles of PRO^{FUR} and PRO^{PC1} suggests that PRO^{PC1} is more stable with regards to pH dependent unfolding when compared with PRO^{FUR}. Under similar conditions, PRO^{SUB} and PRO^{AQU} are stable with minor changes in conformation. Due to its intrinsically unstructured state, PRO^{SUB} would not be expected to undergo conformational changes as a function of pH. However, studies have suggested that an increase in proton concentrations can induce molten-globule like states into unfolded proteins [141-144]. Our studies suggest that acid induced folding is not observed in case of PRO^{SUB}. It is noteworthy that the pH-associated structural transitions PRO^{PC1} and PRO^{FUR} correlates with organelle-specific pHs necessary for activating the mature catalytic domains, MAT^{PC1} and MAT^{FUR} [88]. We next investigated whether propeptides alone are sufficient for pH-dependent activation of cognate proteases, *in vitro* and in tissue culture cells.

Swapping propeptides between PC1 and furin reassigns pH-dependent activation

To monitor *in vitro* activation of propeptide: protease inhibition complexes, we measured enzyme activity as a function of pH (see Methods). Fig. 1.2D demonstrates that PRO^{FUR}:MAT^{FUR} and PRO^{PC1}:MAT^{PC1} show maximum activation at pH~6.5 and pH~5.5, respectively, consistent with the optimal

activation pH of their zymogens [88]. However, the PRO^{PC1}:MAT^{FUR} complex (wherein PRO^{PC1} substitutes PRO^{FUR}) forces the catalytic-domain of furin (MAT^{FUR}) to now display PC1-like activation. Similarly, replacing PRO^{PC1} with PRO^{FUR} causes the catalytic-domain, MAT^{PC1}, to alter its activation to mimic furin (pH~6.5; Fig. 2.2D). Together, the CD spectroscopy, sequence/structural congruence with PRO^{SUB}, and the reassignment of activation pH by swapping PRO^{FUR} and PRO^{PC1} support the hypothesis that eukaryotic propeptides recognize and regulate pH-dependent activation of their cognate proteases *in vitro*.

PRO^{PC1} and PRO^{FUR} control folding and activation of MAT^{FUR} and MAT^{PC1} in cells

Since only correctly folded secretory proteins are efficiently transported from the ER [145], we measured catalytic activities of chimeras (PRO^{PC1}:MAT^{FUR}; ~78kDa and PRO^{FUR}:MAT^{PC1}; ~66kDa; Fig. 1.3A) as readouts for folding/activation, using wild-type constructs (PRO^{FUR}:MAT^{FUR} and PRO^{PC1}:MAT^{PC1}) as controls. All constructs display protease activity when compared with mock transfections (Fig. 2.3B). To monitor primary cleavage of propeptides, we inserted FLAG epitopes between the C-termini of propeptides and N-termini of proteases. These epitopes, which do not affect trafficking, activation, or activity of furin [89, 146], confirm presence of the processed MAT^{FUR} (~69kDa) and MAT^{PC1} (~57kDa) in the conditioned media when probed using Western Blot analysis (Fig. 2.3C, top panel as indicated by arrowheads). Together, the catalytic activities and western blots establish that propeptides can

assist folding and activation of their paralogues in cells. To isolate cleaved inhibition complexes, we assayed each variant using ER-localized constructs wherein the transmembrane- and cytosolic-domains were replaced by the ER localization motif -Lys-Asp-Glu-Leu (KDEL) (Fig. 2.3A). The KDEL motif restricts zymogen reporters to the neutral pH environment of the ER [147], where constructs undergo primary propeptide-cleavages that form PRO:MAT complexes but are blocked from secondary cleavages [94]. Western blot analyses (see Methods) confirm the presence of both unprocessed and processed precursors. Fig. 2.3C (lower panel as indicated by arrowheads) demonstrates the presence of both the protease domain that has undergone the primary processing step to generate the 69kD and 57kD forms of furin and PC1, respectively, as well as the immature protease that has yet to undergo this processing step (78kD and 66kD for furin and PC-1, respectively). The constructs expressing the furin protease domain reliably express at higher levels in the cell culture system we have chosen, regardless of which propeptide it is in complex with, thus the ease in visualizing the two different species. In contrast, the PC1 protease does not express as strongly. Nonetheless there is evidence of both the 66kD unprocessed and 57kD processed forms of PRO^{FUR}-MAT^{PC1} as indicated by the arrowheads in Fig. 2.3C, although the efficiency of this processing is less than the PRO^{PC1}-MAT^{PC1} which appears to undergo efficient autoprocessing under similar conditions. The differences in the processing efficiency of furin and PC1 may reflect differences in the catalytic domains.

We next examined pH-dependent activation of these KDEL-tagged chimeras using wild-type KDEL-tagged reporters and mock transfections as positive and negative controls, respectively. Fig. 2.3D confirms that while maximal activation of PRO^{FUR}:MAT^{FUR}-KDEL occurs at pH ~6.5, the activation of PRO^{PC1}:MAT^{FUR}-KDEL shifts to pH ~5.5. Conversely, activation of PRO^{PC1}:MAT^{PC1}-KDEL shifts from pH~5.5 to pH~6.5 when PRO^{FUR} is used to fold MAT^{PC1}-KDEL in COS-7 cells. Although experiments conducted using crude membrane fractions have higher background activity, they nonetheless confirm that propeptide-dictated reassignment of pH-dependent activation is consistent with the *in vitro* activation of inhibition complexes (Fig. 2.2D) and demonstrates that propeptides regulate compartment-specific pH-dependent activation of furin and PC1.

Histidine-protonation alters conformational dynamics of eukaryotic propeptides

Based on experimental studies, we had hypothesized that the protonation of His^{69} and potentially other histidines may induce conformational changes within PRO^{FUR} to mediate pH dependent activation [103]. Moreover, although His^{69} is conserved, PRO^{PC1} undergoes its pH dependent activation at a much lower pH (5.0). To better understand how histidine protonation may influence propeptide conformations, we conducted MD simulations on PRO^{FUR} and PRO^{PC1} with unprotonated (pH 7) or protonated (pH 6) histidine residues, using PRO^{SUB} and PRO^{AQU} from prokaryotes as controls. MD simulations[148] can provide information that complements biophysical and biochemical studies on

mechanisms of propeptide-mediated protease activation in eukaryotes (see Methods). Early MD simulations of the unfolding of reduced bovine pancreatic trypsin inhibitor (BPTI) on a 500ps time scale suggest the formation of a molten-globule like state that was compact but expanded relative to the native BPTI (11-25%) that is consistent with experimental data [149, 150]. MD simulations have also analyzed the structure and fluctuations of "native" apomyoglobin in aqueous solution for a period of greater than 0.5 nanoseconds and has yielded a detailed model for structure and fluctuations in apomyoglobin which complements the experimental studies [151]. Unfolding simulations using MD methods have yielded insights into the mechanism of extreme unfolding cooperativity in the kinetically stable alpha-lytic protease, a protein that exploits the mechanism of propeptide-dependent folding [152]. In these studies the simulated alpha-lytic protease unfolding pathway produces a robust transition state ensemble that is observed within the 10ns simulation and is consistent with prior biochemical experiments demonstrating that unfolding proceeds through a preferential disruption of the domain interface. Furthermore, the authors demonstrate that αLP unfolds extremely cooperatively while, trypsin, a protein that folds independent of its propeptide, undergoes gradual unfolding under identical conditions of simulations. MD simulations studies have also been used to investigate the role of hydrogen bonding involving the backbone in hen egg white lysozyme, using native as well as partly and fully thionated lysozyme [153]. The results of the simulations show that the structural properties of fully thionated lysozyme clearly differ from those of the native protein, while for partly thionated

lysozyme changes only slightly when compared to native lysozyme. In these studies, the extent of observed unfolding remains constant after 10ns. Hence in our studies are performed MD simulations on a 10ns time-scale. We compared the similarity of structures to the starting conformation by measuring the root-mean-square deviation (RMSD) values at alpha carbons in every residue of the propeptide-domain, along equally spaced snapshots of the simulation trajectory. Simulations suggest that while PRO^{SUB} and PRO^{AQU} are stable, PRO^{PC1} and PRO^{FUR} display enhanced conformational dynamics (Fig. 2.4A and B). Our time-evolved, pH-dependent, residue-specific conformational dynamics suggest that although eukaryotic propeptides display local fluctuations at neutral pH, histidine protonation enhances overall movement and potentially exposes the compartment-specific second cleavage-site loop for proteolysis in PRO^{FUR} (residues 70 to 80) when compared with PRO^{PC1}, which is more stable at pH~6.0-7.0 (Fig. 2.4A). Under identical conditions, PRO^{SUB} and PRO^{AQU} from prokaryotes display remarkable stability towards histidine protonation (Fig 1.4B). To further dissect the structural changes, we plotted the global unfolding of the PRO^{FUR} and PRO^{PC1} as a function of time and at the two different pHs (Fig 2.4C). Global unfolding (Qscore), which was computed using the fraction of native contacts that are retained as a function of time during the simulation at different pHs, demonstrates that PRO^{FUR} appears to undergo significant changes in the native-like contacts upon protonation of the histidine residues. Under similar conditions, PRO^{PC1} appears to be more stable at both pHs.

Since our model for PRO^{FUR} is based on a homology model derived from the NMR structure of PRO^{PC1}, it can be argued that the model may not correspond to an energetically favorable conformation and the simulations may be biased by the homology model. To address this issue we have performed two additional independent simulations on PRO^{FUR} and PRO^{PC1} and for a longer time scale (Fig. 2.4D). To analyze the structural changes we plotted the RMSD of the core and the secondary cleavage site loop between the initial structure and equally spaced snapshots of the trajectory of simulation, both as a function of time and at two different pHs (Fig 2.4D). While PRO^{PC1} remained stable at both pHs, PRO^{FUR} showed increasing RMSD values throughout the simulation at pH 6, while remaining stable at pH 7. The results confirm our earlier simulations on a shorter time scale and suggest that protonation/deprotonation of histidines play a role in the conformational destabilization of PRO^{FUR} compared to PRO^{PC1}. While our simulations do not provide information on why PRO^{PC1} is more stable that PRO^{FUR} towards pH dependent unfolding, they corroborate our experimental observations on the pH dependent stabilities of the propeptides. His^{69} in furin and the corresponding His residue in PC1 reside closely to other histidines and charged residues in the cleavage loop (Fig. 2.4E). The interaction of this protonated His with these other residues may provide key insights into why the activation pHs of furin and PC1 differ dramatically.

Together with our biophysical, biochemical and cell-based studies, the MD simulations suggest that upon protonation of His residues, PRO^{FUR} undergoes conformational changes that may potentially destabilize the propeptide domain to

expose the internal cleavage site for proteolysis. Given that PRO^{PC1} undergoes activation at pH ~5.5 in the DCSGs and remains stable upon His protonation, we can conclude that either additional residues must play a role in the activation of PRO^{PC}, or the timescale of the simulations is too short to capture the unfolding event.

Discussion

Compartmentalizing metabolic pathways within organelles enables eukaryotic cells to process numerous spatiotemporal reactions with efficiency and precision. Optimal organelle function requires maintenance of luminal-pH and propeptide-dependent eukaryotic proteases must have evolved from prokaryotic orthologs to exploit this proton gradient as energy currency to function only at appropriate sub-cellular compartments [97, 136]. Although structures and functions of individual protein families may impose unique evolutionary constraints, an analysis of divergence patterns suggests that individual responses of most proteins are variations on a common set of selective constraints [137]. In protein families with low divergence, mutations within the interior are limited by strong evolutionary pressures to maintain a conserved core that removes all but a few conservative changes [136]. With increasing divergence, mutations in the interior become more widespread and closer in number to what is found in the intermediate and exposed regions [154]. Since catalytic domains exhibit a higher degree of conservation within subtilases when compared to their propeptides [122], this suggests that the catalytic and

propeptide domains may have encountered different mutational frequencies and different selective constraints.

Our work provides insight as to why nature may have imposed differential selective constraints that alter both sequence and the asymmetrical distribution of histidines in two functional domains, namely the propeptides and their cognate catalytic domains within furin and PC1. In this manuscript we demonstrate that PRO^{FUR} and PRO^{PC1} are enriched in histidine-content when compared with cognate proteases and prokaryotic orthologs (Fig 2.1C). Spectroscopic studies demonstrate that changes in pH can induce conformational changes only within PRO^{FUR} and PRO^{PC1}, while their prokaryotic orthologs, PRO^{SUB} and PRO^{AQU}, are largely unaffected (Fig 2.2A and B). Since swapping propeptides between eukaryotic paralogues transfers pH-dependent protease activation in an propeptide-dictated manner (Fig. 2.2D and 2.3D), while allowing folding and cellular localization (Fig. 2.3B and C), our results argue that PRO^{FUR} and PRO^{PC1} may have evolved from prokaryotic orthologs to encode histidine-driven pH-sensors that enable furin and PC1 to recognize and adapt to cellular organelles. Our MD simulations suggest that histidine protonation may be sufficient to induce conformational changes that enable the second activating cleavage of the propeptide and are consistent with our spectroscopic analysis (Fig. 2.2B and C). While it would be interesting to compare the structures of the chimeras with those of the wild-type complexes and examine how their structures are affected by changes in pH, such experimentation is currently unfeasible due to the high concentrations of protein required for circular dichroism spectroscopic analysis.

It is important to note that despite histidine enrichment, the specific location of these residues within the amino acid sequences of propeptides can vary significantly (Fig 2.1A and 2.4E). Moreover, the His^{69} that was identified as a primary pH sensor in PRO^{FUR} [103] is also conserved in PRO^{PC1}, although the pH dependent activation of furin and PC1 differs significantly [88]. This suggests that additional undetermined residues and/or cellular factors must play a significant role in pH dependent activation of their cognate protease domains. Propeptides also contain several charged residues [5] which may interact with protonated and non-protonated histidine residues, thereby enabling subtleties in their sensitivity to compartment specific pH. Hence, our studies emphasize the necessity of more detailed analyses of the differences between pH-sensors of PRO^{FUR} and PRO^{PC1} using detailed site-directed mutagenesis studies, to tease out the interplay with residues in the proximity of their cognate pH sensors.

Since propeptides facilitate the folding of several eukaryotic proteases, this raises the possibility that other propeptide dependent eukaryotic proteases may also display similar bias towards His-residues. Cathepsins represent another example where preliminary results suggest similar histidine enrichment within propeptides (Elferich, unpublished data). Cathepsins also undergo compartment specific activation of their cognate catalytic domains within the acidic pH of the lysosomes (pH 4.0). However, unlike furin and PC1, which can be compared with prokaryotic orthologs from the ubiquitously expressed subtilase super-family, cathepsins do not have well characterized prokaryotic orthologs to precisely compare histidine enrichment as a function of prokaryotic

versus eukaryotic evolution. Interestingly, the histidine residues localized within the propeptides are likely to modulate a wide range of pH dependent activation. Hence, at low pH typically found within the lysosome, all of the histidine residues are likely to be protonated if their pKa is not altered by their structural context. Therefore it is possible that other residues such as aspartic and glutamic acid residues may collaborate with histidines to mediate subtle changes in pH dependent activation. Other residues could either become protonated themselves to mediate activation or influence the pKa of histidine protonation. It is also possible that the pH dependence in activity for furin and PC1 could also partially reflect the pKa values of catalytic residues and would required detailed characterization of active site residues. The challenge is to understand which specific histidines interact with additional residues to provide a broad range of pH-dependent activation of secretory proteases.

Materials and methods

Expression and purification of PRO^{FUR}, PRO^{PC1}, MAT^{FUR} and MAT^{PC1}:

Codon optimized genes encoding human PROFUR and mouse PROPC1 were synthesized from CELTEK genes, cloned into pET11a and expressed in BL21(DE3) as described [76]. Inclusion bodies containing MAT^{FUR} and MATPC1 were isolated and proteins were purified using reverse phase chromatography. Enzymatically active MAT^{FUR} and MAT^{PC1} were obtained from recombinants expressing human VV:fur/f/ha/ΔTCK[103] and mouse VV:mPC1 [155] in BSC40 cells as described [103]. Cos7 cells were maintained in DMEM-high glucose

medium (HyClone) containing 10% fetal bovine serum and 1% penicillin-streptomycin. Cells were incubated at 37oC in a 5% CO2 environment as described [103].

Circular dichroism studies:

Circular dichroism (CD) measurements were performed on an AVIV model 215 CD spectrometer using a 1 mm path-length cell at 4 oC as described earlier[65, 66]. Briefly, propeptide samples (4 mg/ml) stored in 6M GdnHCl (to avoid side-chain modifications commonly seen when samples are stored in urea) were diluted to a final concentration of 0.4 mg/ml), and were refolded using stepwise dialysis against 50 mM cacodylate buffer, pH 7.4 containing 150 mM KCl (Buffer A) and decreasing amounts of urea. The proteins were dialyzed twice in Buffer A without urea, against Buffer A in different pH (5.0-7.0), and then subjected to ultracentrifugation in TLA-100 for 30 min to remove particulates. The CD spectra between 200-260 nm were averaged over three independent experiments and plotted as a change in ellipticity at 222 nm as a function of pH and plotted as [q] molar ellipticity[156] deg.cm2.dmol-1. The PROSUB-C structure was obtained by a difference spectra between the cleaved PROSUB:S221C-subtilisin complex and mature subtilisin as described earlier[42].

Molecular Dynamic simulations:

1SCJ[47], 1KN6[91] and homology models of furin derived from 1KN6, and aqualysin derived from 1SCJ were used for as PDB models PRO^{SUB}, PRO^{PC1}, PRO^{FUR}, and PRO^{AQU} respectively. Homology models were built using

either SWISS-MODEL or MODELLER. All hydrogen and non-protein atoms were removed and hydrogen were added back using the autoPSF function in NAMD [157]. Structures were solvated in cubes with TIP3P explicit water using VMD, with a minimum of 12 Å distance to the edge. All simulations were carried out with periodic boundary conditions, PME for long-range electrostatics, and a 12 Å cutoff for non-bonded interactions with the CHARMM22 force field[158] using NAMD (version 2.5). Snapshots were saved every 10 ps using a time-step of 1 fs. The system was equilibrated by first constraining the protein and minimizing solvent for 1000 steps using a conjugate gradient algorithm. The solvent was initially equilibrated for 100 ps, then fully constrained and the protein minimized for 500 steps. The entire system was subsequently minimized and used in the simulations. MD simulations require defining of a potential function or a force field that describes the ways through which particles in a simulation will interact[159]. Force fields can be defined at many levels of physical accuracy and those used in MD-simulations often embody a classical treatment of particle-particle interactions, which can reproduce structural and conformational changes, but usually cannot reproduce precise chemical reactions. Therefore, to simulate the pH-dependent protonation reactions, we have approximated the pH environment by predetermining the protonation state in the starting structure, an approach that has been extensively employed in the field of molecular dynamics. For pH 7, we used the HSD parameters for histidine residues which represent an uncharged side chain, with a proton bound to the nitrogen atom in the delta position. To simulate an environment of pH 6 we used the HSP parameter which represents a

positively charged histidine with protons bound to both nitrogen atoms. For testing the robustness of our simulations we took two different models of PRO^{FUR} and PRO^{PC1} and repeated the simulations as described above. An adjustment of the pH to exact values would require a prediction of the pKa values of individual residues, which was not practical in the given study.

Amino acid content analysis:

Protein sequences for human furin, mouse PC1, subtilisin from *Bacillus subtilis*, and aqualysin from *Thermus aquaticus* families were obtained from the 50% sequence identity clusters UniRef50_P09958, UniRef50_P29120, UniRef50_P00782, and UniRef50_P08594 in the UniRef database, respectively. Subsequences representing the propeptides and the protease domain were extracted using annotation from the Interpro database entries IPR009020 and IPR000209, respectively. Sequences that were not annotated by both entries were omitted. Amino acid content of both domains in all sequences were calculated and averaged for each domain and protein family. Contents of amino acids belonging to individual groups were added and divided by the sum of their content in the whole UniProt database (release 2011_12). The multiple sequence alignment of selected prokaryotic and eukaryotic subtilases was obtained using ClustalW and colored using Genedoc.

Secreted enzyme activity assays:

For all assays, 113 μM furin substrate (Abz-RVKRGLA-Tyr[3-NO_2]) in dimethyl sulfoxide was incubated with 40 μl secreted enzyme in 155 μl of 50 mM cacodylate buffer, pH 7.0 containing 1 mM $CaCl_2$ and 50 mM KCl. Cacodylate

buffer was used in all experiments to maintain consistency throughout the analyses. The assays were conducted on a SpectraMax-M2 spectrofluorometer equipped with a 96-well plate reader. Excitation wavelength was set at 320 nm while emission wavelength was set at 425 nm. Given values are averages of triplicate assays. The activity was normalized by quantifying the relative amounts of proteins secreted in the media using ImageJ software.

Isolation of in trans propeptide: protease complexes:

Since propeptides are potent competitive inhibitors of protease paralogues[76], PRO:MAT complexes *in trans* were generated by adding 10-fold excess of PRO^{FUR} and PRO^{PC1} (~2 nM) to MAT^{FUR} or MAT^{PC1} (~0.2 nM) in 50 mM cacodylate buffer, at different pH (5.0 to 7.4) containing 150 mM KCl in a 96-well quartz plate. Complexes were incubated for 30 min at RT and the activities assayed as described earlier[103]. The percent activity at each pH was calculated using the activity of uninhibited protease as a control.

Construction of secreted and ER localized PCDNA3.1 expression vectors:

The plasmid p2Vneo containing human furin[89] was cut with EcoRI and HindIII to release the Furin-Flag gene. The plasmid pBSSK containing the mouse PC1-Flag gene[155] was cut with NcoI and BamHI. Both genes were treated with Klenow and ligated with PCDNA3.1 cut using EcoRV. Gene orientations were confirmed by digestion with BamHI and XhoI and through DNA sequencing. To obtain soluble and secreted PRO^{FUR}-Flag-MAT^{FUR}, a stop codon was introduced after Leu^{713} in the plasmid containing full-length furin-flag, which removes the

cysteine-rich, cytoplasmic, and transmembrane domains. Similarly, the PC1-Flag plasmid was truncated at Arg^{618} to produce secreted PRO^{PC1}-Flag-MAT^{PC1}. The chimeras (PRO^{FUR}-Flag-MAT^{FUR}, PRO^{FUR}-Flag-MAT^{PC1}, PRO^{PC1}-Flag-MAT^{PC1}, PRO^{PC1}-Flag-MAT^{FUR}) were constructed using PCR. To localize proteases in ER, KDEL sequences were inserted into the genes expressing soluble PRO^{FUR}-Flag-MAT^{FUR} and PRO^{PC1}-Flag-MAT^{PC1} using PCR. All constructs were confirmed through sequencing (OHSU DNA Services Core, Portland, OR).

Expression of constructs and ER fractionation:

For secretion experiments, cells were transfected with expression vectors containing PRO^{FUR}-Flag-MAT^{FUR}, PRO^{FUR}-Flag-MAT^{PC1}, PRO^{PC1}-Flag-MAT^{PC1}, PRO^{PC1}-Flag-MAT^{FUR} or the empty PCDNA3.1 vector using LT1 transfection reagent (Invitrogen) as recommended by the manufacturer. After 5 hrs, the cells were washed with PBS, replaced with serum free DMEM media and CM was harvested 24 hrs post media change. For the ER retention experiments the KDEL-tagged constructs were transfected as described above. The media was not changed after transfection and the cells were maintained in 10 cm plates. The microsomal fraction was prepared following manufacturer's instructions from Endoplasmic Reticulum Isolation Kit (Sigma). Constructs were probed by Western blotting as follows: primary antibody, mAB M2 flag (Sigma), was used in a 1:1000 dilution and the secondary antibody, IgG3000 anti-mouse (Fisher), was used in a 1:10,000 dilution.

KDEL enzyme activity assays:

Cells were transfected with constructs containing PRO^{FUR}-Flag-MAT^{FUR}-KDEL, PRO^{FUR}-Flag-MAT^{PC1}-KDEL, PRO^{PC1}-Flag-MAT^{PC1}-KDEL, PRO^{PC1}-Flag-MAT^{FUR}-KDEL[103]. After 24 hrs, harvested cells were lysed and incubated in a 25 ^{o}C water bath for 1 hr in 50 mM cacodylate buffer, pH 7.4 containing 1 mM $CaCl_2$, 150 mM KCl, and a fresh protease inhibitor cocktail (Sigma). To six micro centrifuge tubes containing 95 µL of 100 mM cacodylate buffer of varying pH, 100 µL of cell lysate was added to bring the mixtures to a final pH of 5.4, 6.4, and 7.4, in duplicates. The cell lysates containing processed enzymes were incubated at 25^{o}C for 2 hrs with one tube from each pH incubated with 0.83 nM trypsin (as a control for complete activation). Soybean trypsin inhibitor was added to the tubes (to block trypsin, which can interfere with the activity assay) and they were incubated for an additional 15 minutes. For each assay, 113 µM furin substrate was added for a final volume of 200 µL. The assays were conducted as described earlier. Each experiment was performed at least 3-times and the values given are the average of assays done in triplicate.

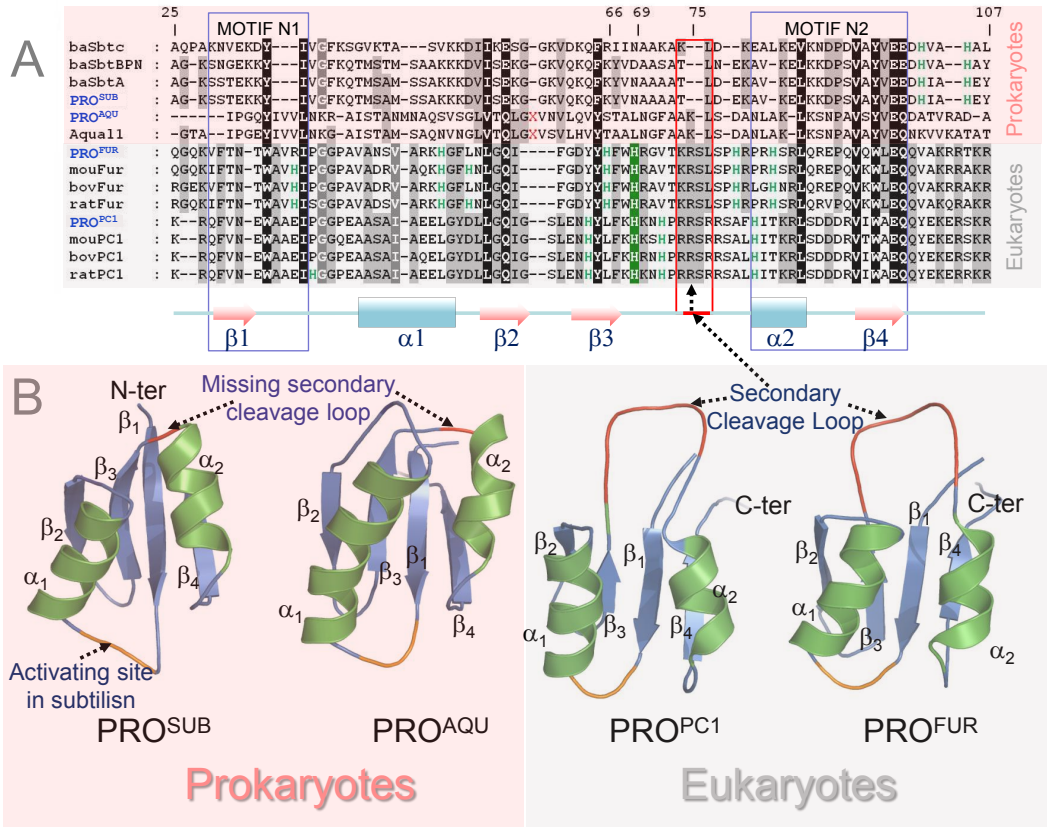
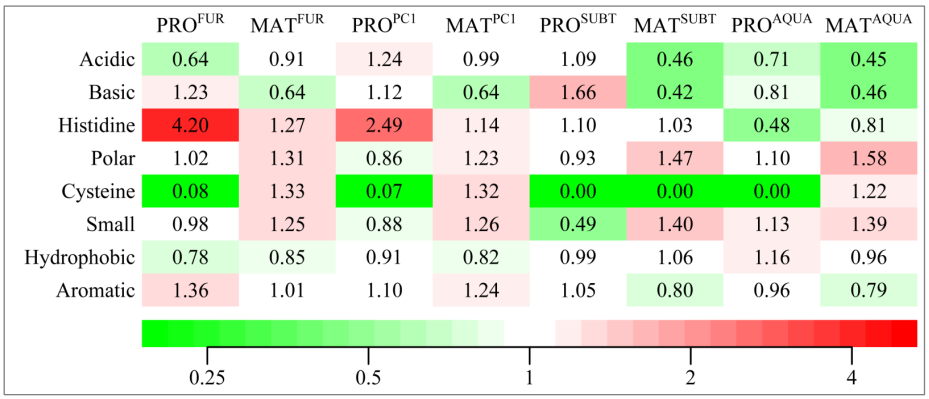


Figure 2.1: Comparison of sequences, structures, evolution and composition biases of propeptides in prokaryotic and eukaryotic subtilases.

The pink and grey background in panel A through D indicates prokaryotes and eukaryotes, respectively. **(A)** Multiple Sequence Alignment (MSA) displaying conservation between eukaryotic subtilases and prokaryotic orthologs. Numbering is based on furin. Residues shaded black are 100% conserved, dark grey >80%, and light grey >50%. The conserved pH-sensor in furin is shaded green and the secondary cleavage loop is indicated by the red box. Red X's represents an insertion of 5 residues in aqualysin. Pink shading represents prokaryotes while the light gray represents eukaryotes. Secondary structures displayed below MSA are based on PRO^{PC1} (1KN6). Motifs N1 and N2 depict folding nucleation sites for MAT^{SUB}. **(B)** Structures of propeptides displayed as ribbon diagrams. PRO^{SUB} structure was extracted from the propeptide: subtilisin (1SCJ), while PRO^{AQU} structure is a homology model based on 1SCJ and 2W2M. The structure of PRO^{PC1} is derived from the NMR (1KN6) while PRO^{FUR} represents a homology model of PRO^{PC1}[103].

C



(C) Heat map displaying amino acid content within the propeptides and catalytic domains of prokaryotic subtilisin and aqualysin and eukaryotic PCs, furin and PC1. Protein sequences for furin (n=26), PC1 (n=14), subtilisin (n=69), and aqualysin (n=7) families were obtained from the 50% sequence identity clusters UniRef50_P09958, UniRef50_P29120, UniRef50_P00782, and UniRef50_P08594 in the UniRef database, respectively. Amino acid content for each family of propeptides and protease domains were calculated and averaged. Contents of amino acids belonging to individual groups were added and divided by the sum of their content in the whole UniProt database (release 2011_12) to obtain the fold change. Within an individual group, fold values greater than one indicates residue enrichment, values less than one indicate residue depletion while a value of one indicates no change.

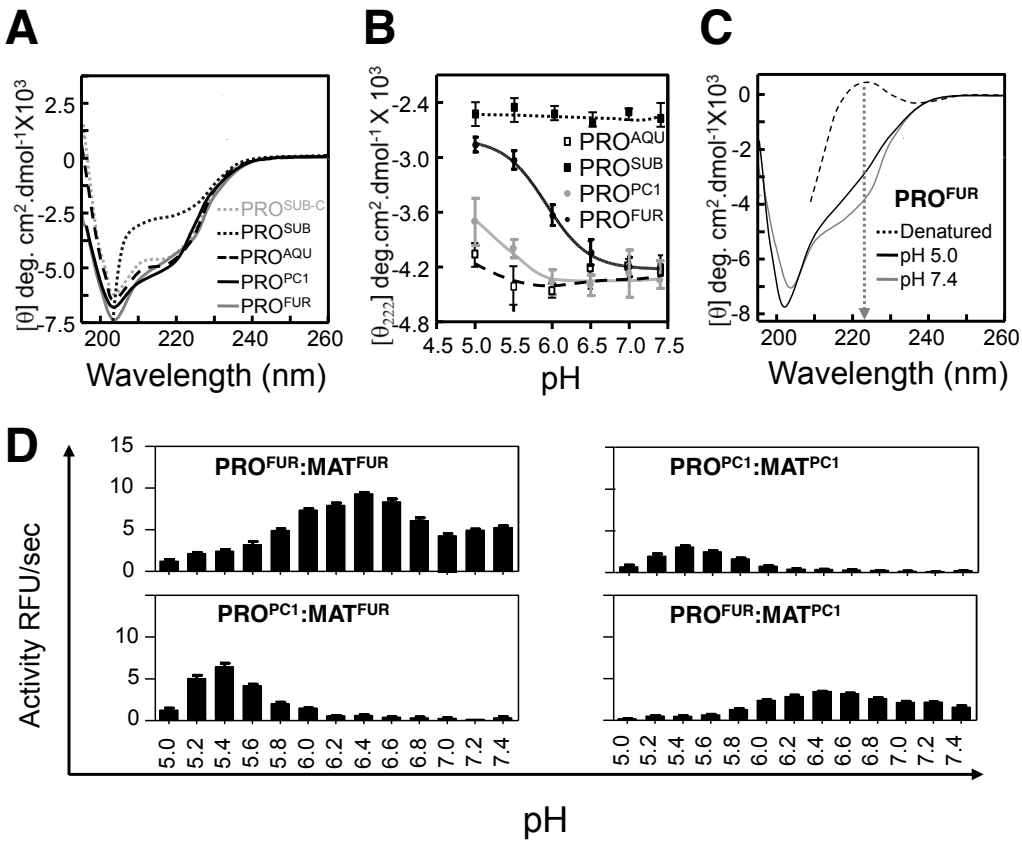


Figure 2.2: pH dependent structure and function propeptides.
(**A**) Secondary structures determined using CD spectroscopy performed at a pH 7.0 and plotted as molar ellipticity [q] deg.cm^2.dmol^{-1}. (**B**) Structural stability of propeptides monitored by changes in ellipticity at 222 nm as a function of pH. (**C**) The secondary structure of PRO^{FUR} at pH 7.4 and 5.0, compared with completely denatured furin. The arrow marks 222 nm on the scale. (**D**) Type of eukaryotic propeptide dictates pH-optimum for activation of propeptide: protease complex. The activation optimum for MAT^{FUR} shifts from pH~6.5 in the presence of PRO^{FUR} to pH~5.5 when PRO^{PC1} forms the complex. Conversely, MAT^{PC1} activation shifts from pH~5.5 in presence of PRO^{PC1} to pH~6.5 when PRO^{FUR} forms its complex. Values are measurements of three independent experiments.

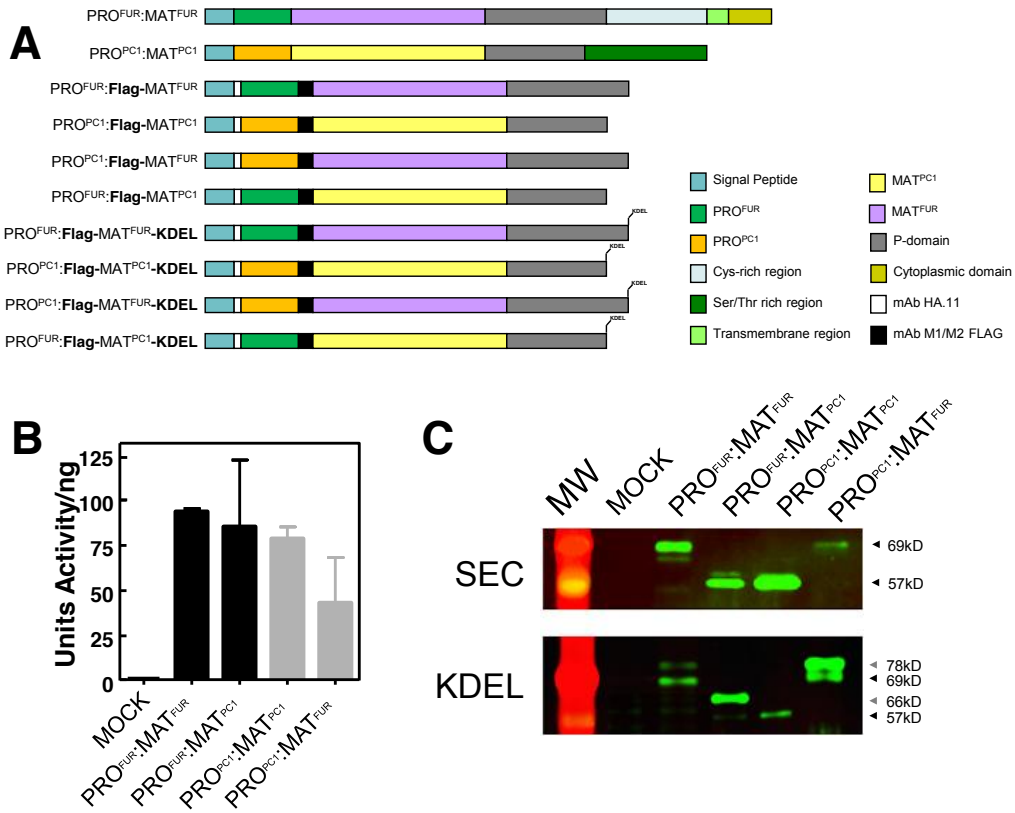
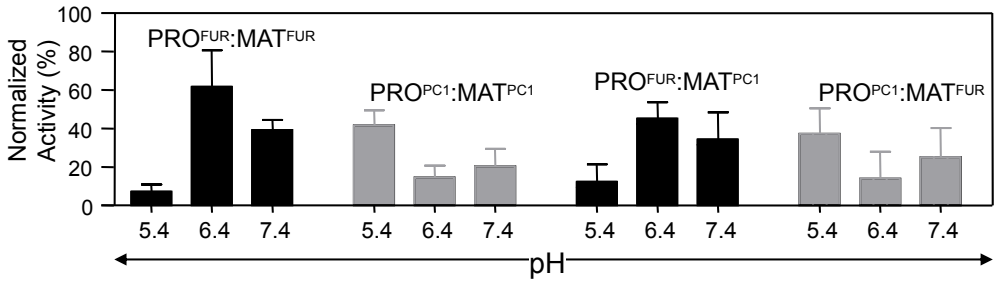


Figure 2.3: Transferring propeptides between eukaryotic subtilases reassigns their optimum pH for activation.
(A) Schematic of constructs for Furin and PC1. **(B)** Normalized protease activity assayed in conditioned media (CM) from Cos-7 cells transfected with 2 µg of DNA **(C)** Western blot analysis of CM from cells expressing secreted reporter constructs (top panel; SEC), and ER fractions from cells expressing KDEL-tagged reporters probed using mAB-M2. Molecular weight of each species is indicated by the arrowheads; Unprocessed furin, 78kD; Processed furin, 69kD; Unprocessed PC1, 66kD; Processed PC1 57kD.

D



(D) pH-dependent activation of KDEL-tagged reporters measured after incubating ER membrane fractions at designated pH[103]. Maximal activity was estimated by trypsinizing membrane fractions for 1 hr and inhibiting trypsin by soybean trypsin inhibitor prior to the protease assay.

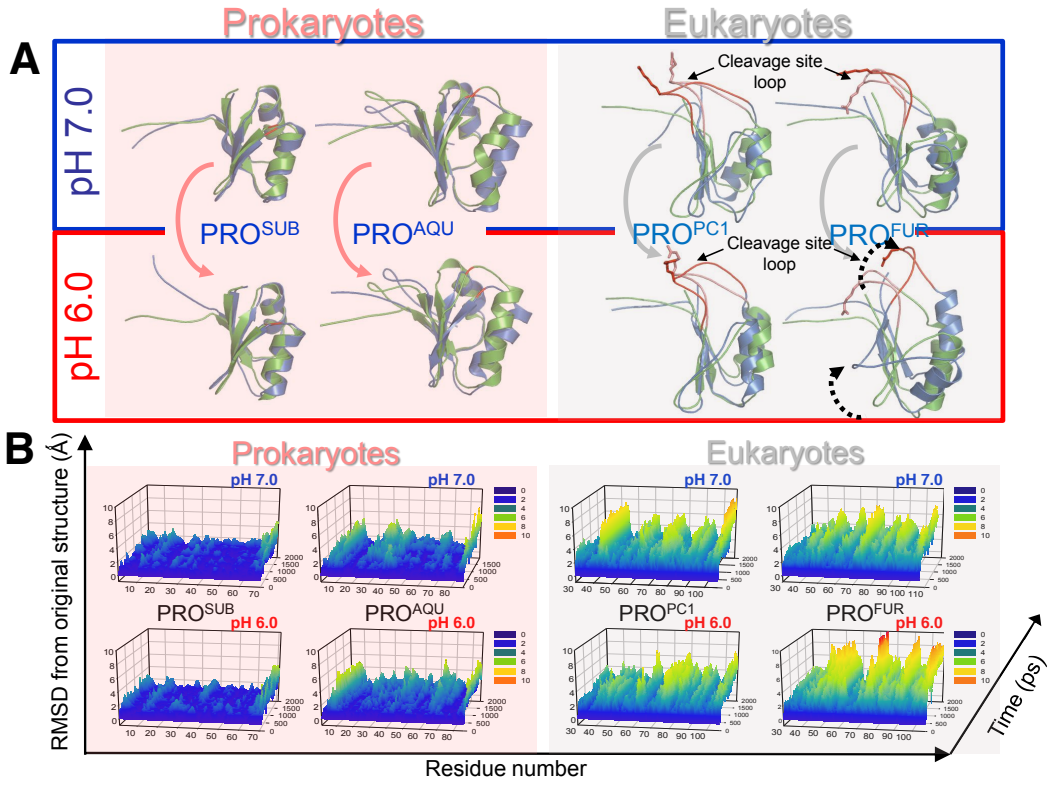
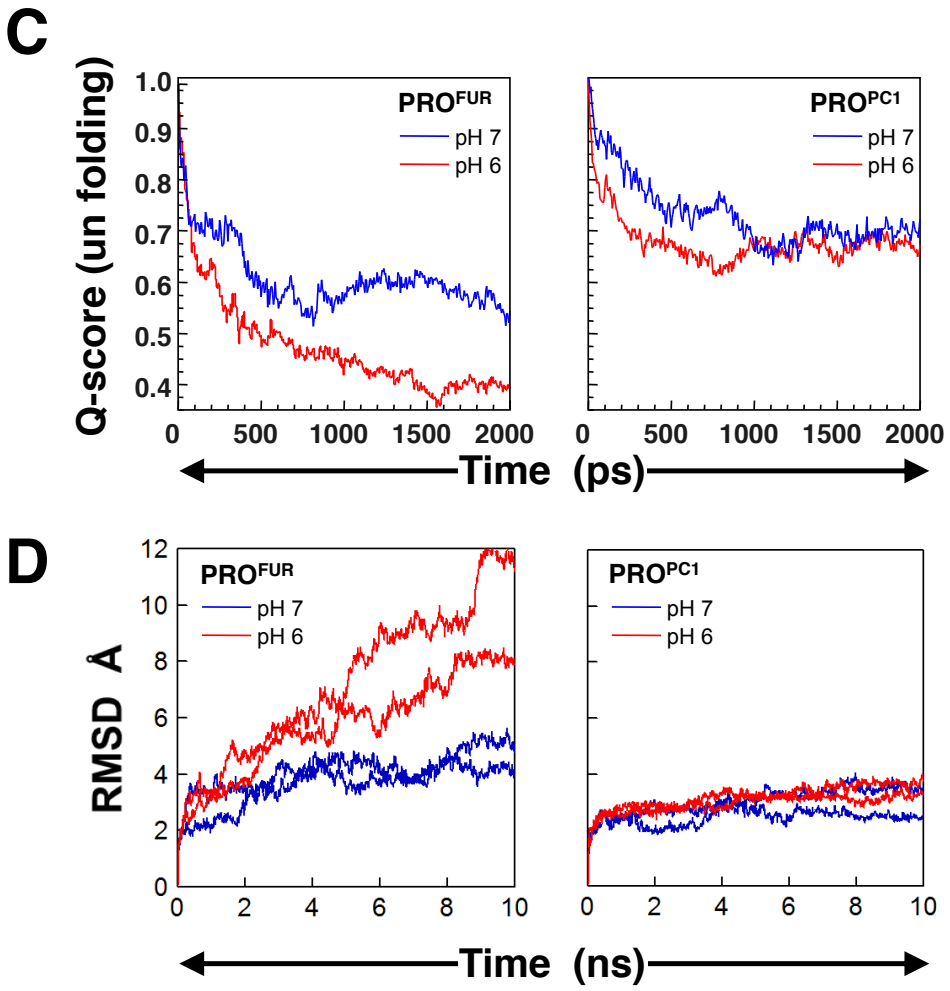
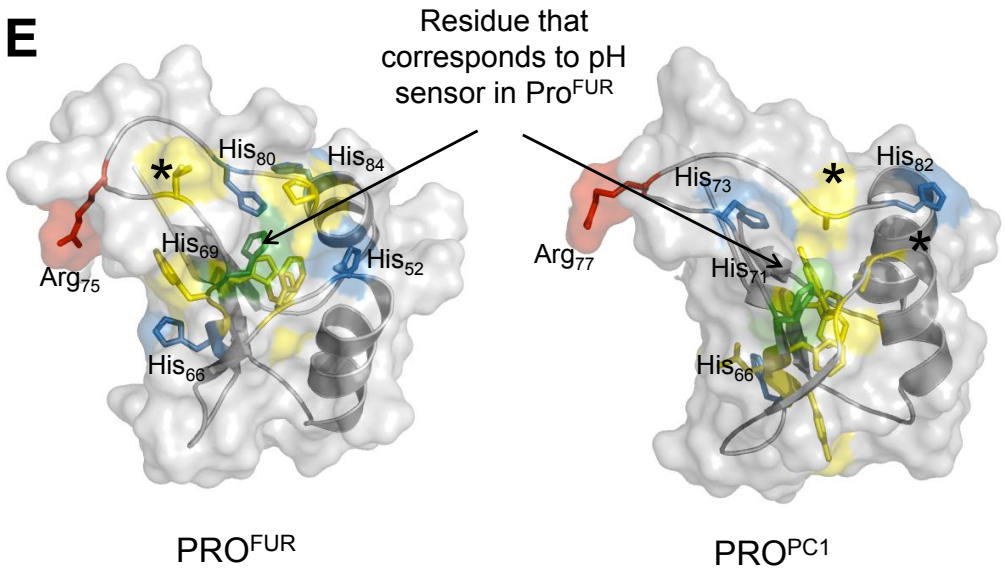


Figure 2.4: pH dependent structural dynamics of prokaryotic and eukaryotic propeptides.
(**A)** Green and blue cartoons represent initial and final structures of the simulations, respectively. The second cleavage site loop (red/salmon) in PRO^{FUR} (structure on right-side) is stable when histidines are deprotonated (pH 7.0; bordered by black box) but changes conformation upon histidine-protonation (pH 6.0; bordered by red box). The dynamics of the loop are unaffected by the histidine-protonation status of PRO^{PC1} (cartoons on left-side). Under identical conditions PRO^{SUB} and PRO^{AQU} show insignificant changes in dynamics as a function of pH **(B)** Protonation status dependent, time-resolved, residue-specific dynamics of PRO^{SUB}, PRO^{AQU}, PRO^{PC1} and PRO^{FUR}. Arrowhead indicates secondary cleavage-site and color scale represents RMSD from initial structures.



(C) Global unfolding (Qscore) of PRO^{FUR} and PRO^{PC1} at different pHs. Unfolding was computed using the fraction of native contacts that are retained as a function of time during the simulation at different pHs and suggest that PRO^{FUR} undergoes global unfolding at a pH of 6.0 when compared with pH 7.0 and with PRO^{PC1} at both, pH 7.0 and 6.0, respectively. **(D)** Evaluating the robustness of independent MD simulations using different models and longer time scales. We compared the similarity of structures to the starting conformation by measuring the root-mean-square deviation (RMSD) within the propeptide-domain, along equally spaced snapshots of the simulation trajectory. Our results suggest that while PRO^{PC1} appears stable at different pHs, PRO^{FUR} displays significantly larger conformational changes, which may contribute to its increased proteolytic susceptibility at pH 6.0, and is consistent with our spectroscopic studies.



(E) A comparison of the structural locations of various histidine residues in PRO^{FUR} and PRO^{PC1}. The pH sensor His_{69} in PRO^{FUR} (green) along with other histidine residues (blue) and their corresponding residues with PRO^{PC1} are depicted. Hydrophobic residues surrounding His^{69} in PRO^{FUR} are depicted in yellow, while the asterisks denote residue substitutions at cognate histidine residues.

Chapter III: The mechanism by which a propeptide-encoded pH-sensor regulates spatiotemporal activation of furin

Danielle M Williamson[1,†], Johannes Elferich[1,†], Parvathy Ramakrishnan[1,2], Gary Thomas[3] and Ujwal Shinde[1,*]

[1]Department of Biochemistry and Molecular Biology, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland OR 97239, USA

[2]Current Address: Dr. Reddy's Laboratories Limited, Survey No. 42, 45 & 46, Bachupally, Hyderabad – 500072, Andhra Pradesh, India

[3]Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, PA 15219, USA

This chapter was originally published in The Journal of Biological Chemistry in June 2013. Experiments were performed by DMW, with JE contributing the MD simulations, and PR performing the original cloning. The manuscript was written by DMW and US

Summary

The proprotein convertase furin requires the pH gradient of the secretory pathway to regulate its multi-step, compartment-specific autocatalytic activation. While His_{69} within the furin prodomain serves as the pH sensor that detects transport of the propeptide: enzyme complex to the trans-Golgi network (TGN) where it promotes cleavage and release of the inhibitory propeptide, a mechanistic understanding of how His_{69} protonation mediates furin activation remains unclear. Here we employ biophysical, biochemical and computational approaches to elucidate the mechanism underlying the pH dependent activation of furin. Structural analyses and binding experiments comparing the wild type furin propeptide with a nonprotonatable His_{69}→Leu mutant that blocks furin activation in vivo revealed protonation of His_{69} reduces both the thermodynamic stability of the propeptide as well as its affinity for furin at pH 6.0. Structural modeling combined with mathematical modeling and Molecular dynamic simulations suggested His_{69} does not directly contribute to the propeptide: enzyme interface but rather triggers movement of a loop region in the propeptide that modulates access to the cleavage site, and thus allows for the tight pH regulation of furin activation. Our work establishes a mechanism by which His_{69} functions as a pH-sensor that regulates compartment-specific furin activation, and provides insights into how other convertases and proteases may regulate their precise spatiotemporal activation.

Introduction

The requirement for single or multiple endoproteolytic cleavages of precursor pro-proteins as they transit the secretory pathway is an evolutionally conserved theme in the biochemistry of biologically active proteins and peptides [160]. Proprotein convertases (PCs) are a family of calcium dependent serine proteases that process inactive proproteins in eukaryotes [115, 161]. PCs include nine endoproteases; furin, PC1/PC3, PC2, PC4, PACE4, PC5/PC6, PC7/ LPC/PC8, SKI/S1P, and NARC-1/PCSK9, and constitute a subfamily within the subtilase super-family [115, 122]. PCs are synthesized on the endoplasmic reticulum (ER) and translocated into the ER lumen where they undergo folding and intramolecular excision of their prodomains to form a propeptide:PC complex that is competent for export to late secretory compartments, but is catalytically inactive because the propeptide masks the catalytic domain from substrate binding [88]. Upon reaching their correct cellular destinations for processing substrates in *trans*, the propeptide:PC complex undergoes activation, usually through auto-proteolytic cleavage of the bound propeptide (for review, see [115, 122]). Although the synthesis of proteases as inactive proenzymes enables cells to spatially and temporally regulate their catalytic activity, the molecular and cellular determinants that modulate activation of PCs remain poorly understood [122].

Furin, the most extensively studied PC, catalyzes proteolytic maturation of a diverse repertoire of growth factors, receptor, and enzyme precursors within

multiple secretory pathway compartments [88, 115, 122]. Consistent with its essential role in homeostasis, mice lacking furin die before embryonic day 11 due to cardiac defects [162, 163]. Additionally, misregulation of furin activation has been associated with cancer invasiveness and metastasis, and susceptibility to viral and parasitic infections (for Review [116]). The furin precursor contains an 83-residue N-terminal propeptide (PRO^{FUR}) that is essential for folding the catalytic domain (MAT^{FUR}) in the ER [87, 93], putatively by stabilizing the late stages of the folding transition state [5]. Due to the nature of its role in facilitating folding, the furin propeptide is considered an intramolecular chaperone (IMCs)[38, 42]. Once folded, the catalytic domain rapidly cleaves propeptide at the consensus furin site $RTKR_{107}$ ($t_{1/2}$ < 10 min.,[94]), which permits export of the inactive PRO^{FUR}:MAT^{FUR} inhibition complex from the ER to late secretory pathway compartments [89, 93, 164, 165]. Upon reaching the mildly acidic *trans*-Golgi network (TGN; pH ~6.0), PRO^{FUR} undergoes a slow second internal proteolytic cleavage at the noncanonical furin site $_{69}HRGVTKR_{75}$ ($t_{1/2}$ < 100 min., [94]), which disrupts the inhibition complex to allow MAT^{FUR} to exert its catalytic activity [89, 94]. The necessity for exposure of PRO^{FUR}:MAT^{FUR} to the acidic pH of the TGN, coupled with the importance of tight regulation of protease activation, argues for the presence of a sensor that recognizes when the complex reaches the correct pH-environment.

We reported that a conserved histidine residue (His_{69}) in PRO^{FUR} acts as a pH-sensor that regulates the compartment-specific activation of pro-furin [103]. While a $His_{69}Leu$ substitution in the furin propeptide ($H_{69}L$-PRO^{FUR}) permitted

correct folding, prodomain excision and export of the propeptide: furin complex from the ER to the TGN/endosomal system, the substitution blocked the pH-dependent second cleavage of MAT^{FUR} at R_{75}, thereby preventing furin activation [103]. Moreover, we recently reported that that PRO^{FUR} and the propeptide of PC1/3 (PRO^{PC1}) are sufficient to regulate the pH dependent activation of their cognate catalytic domains, suggesting a broad role for pH sensing in activation of PCs [92] and other secreted proteases [166]. Indeed, swapping propeptide-domains transfers sensitivity to pH-dependent activation of the protease domain in a propeptide-dependent manner [92].

Although His_{69} protonation is required for furin activation, the precise mechanism by which this pH-sensor mediates activation has remained unclear. Here, we use WT-PRO^{FUR} and the deprotonated state mimic H_{69}L-PRO^{FUR} to explore structure, stability and pH-dependent binding, coupled with mathematical modeling and molecular dynamics, to understand how His_{69} functions as a pH-sensor. Taken together, our work explains the structural and mechanistic basis by which His_{69} regulates compartment specific furin activation, and provides insight into how other PCs may regulate their own activation.

Experimental procedures

Protein production and purification:

Sequences coding for WT-PRO^{FUR} and H_{69}L-PRO^{FUR} variants were cloned into pET11a and expressed in BL21(DE3) *E.coli* as described [92]. Inclusion bodies were isolated and solubilized after cell lysis performed using a French

press apparatus. Protein was purified using a size exclusion column, and dialyzed into 6M guanidinium HCl for long-term storage. Mature furin was over-expressed in BSC40 cells infected with recombinant vaccinia virus encoding human furin, as described [92, 103].

CD Spectroscopy:

Circular dichroism (CD) spectroscopy was performed on an AVIV model 215 CD spectrometer at 4^{o}C as described [65, 66, 92]. Before taking CD measurements, purified protein was refolded in a stepwise fashion by dialysis against refolding buffer [50 mM Tris, 150 mM KCl, 1 mM $CaCl_2$, pH 7.4] with incrementally smaller amounts of urea, with a final dialysis step into cacodylate buffer [50 mM Cacodylic acid, 150 mM KCl, 1 mM $CaCl_2$, pH 7.4]. An additional dialysis step was carried out to adjust pH as necessary. Refolded protein was then subjected to ultracentrifugation for 15 minutes to remove any particulate, and CD spectra between 190-260 nm taken to ensure complete folding. For the protease domain, 0.1 mg/ml of mature furin was dialyzed against the Cacodylate buffer at pH 5.0 to 7.0 and used for analyzing pH stability.

Urea denaturation:

Thermodynamic stabilities of the wild type propeptide (WT-PRO^{FUR}) and H_{69}L-PRO^{FUR} were measured by perturbing the secondary structure using urea and fitting the data to either two-state or three-state unfolding models as described [65]. WT-PRO^{FUR} or H_{69}L-PRO^{FUR} (~3mg/ml) were refolded as above into a buffer of desired pH and urea was added in 0.01M steps to a final

concentration of 6 M urea using a titrator. Changes in absorbance at 222 nm were monitored as a function of denaturant concentration.

Glycerol stability:

WT-PRO^{FUR} or H_{69}L-PRO^{FUR} (~0.4 mg/ml) was refolded as above, then glycerol added to a final concentration ranging from 0-30%. CD spectra between 190-260 nm was taken to assess the effect of glycerol on stability of the isolated propeptide, as described [65]. Approximately 0.1 mg/ml of mature furin was dialyzed against 50 mM Cacodylic acid, 150 mM KCl, 1 mM $CaCl_2$, at pH 7.0 containing 30% glycerol and used for analyzing the effect of glycerol on the protease conformation.

Activity assay:

Activity assays were performed to determine the inhibitory capabilities of the various propeptides as described [103]. 129 μM of the fluorogenic furin substrate (Abz-RVKRGLA-Tyr[3-NO_2] was incubated with serially diluted amounts of either WT-PRO^{FUR} or H_{69}L-PRO^{FUR} (concentrations ranging 0.15-3000 nM) in cacodylate buffer [50 mM cacodylic acid, 150 mM KCl, 1 mM $CaCl_2$]. 0.5 μl secreted furin was added to initiate the reaction. To assess pH dependence of IC_{50}, the pH of the cacodylate buffer was varied between pH 5.0-7.4. For assays to measure the effect of glycerol on stability, 30% was added or not to cacodylate buffer at pH 7.0. All activity assays were conducted in triplicate on a SpectraMax-M2 spectrofluorometer equipped with a 96-well plate reader, with excitation wavelength set to 320 nm and emission wavelength set to 425 nm

[89]. Data were fitted and analyzed using GraphPad Prism to determine K_i and IC_{50} values.

Molecular Dynamics simulations:

A homology model of WT-PRO^{FUR} derived from the solution structure of the paralogues prohormone convertase 1 propeptide (1KN6) [91], was used as a model of PRO^{FUR}, as described earlier [92]. The model for WT-PRO^{FUR} was obtained by mutating residue histidine at position 69 to a leucine using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.) and minimized using the CHARMM22 force field [158]. All hydrogen and non-protein atoms were removed and hydrogens added back using the autoPSF module in NAMD version 2.5 [157]. Structures were explicitly solvated in water cubes in VMD with a minimum distance of 12 Å to the edge of the cube. All simulations were carried out with periodic boundary conditions, particle mesh Ewald for long-range electrostatics, and a 12 Å cut-off for non-bonded interactions within the CHARMM22 force field, using NAMD. Simulations were run with a step size of 1 fs, with snapshots taken every 10ps, as described earlier [92]. To first equilibrate the system, the protein was constrained, and solvent minimized for 1000 steps using a conjugate gradient algorithm. The minimized system was then used for subsequent simulations. To simulate the pH-dependent protonation reactions, we approximated the pH environment by predetermining the protonation state of the histidines in the starting structure, as described earlier [92]. For a pH environment of 7.0, we used the HSD parameter, which represents the uncharged state, with a proton bound to the nitrogen in the

delta position; for a pH of 6.0, we used the HSP parameter to simulate the positively charged, protonated histidine with protons bound to both nitrogen atoms.

Mathematical modeling:

Modeling of the kinetics of the furin activation step was done using CellWare version 2.0 (Systems Biology Group at the Bioinformatics Institute, Singapore)[167]. CellWare is an integrated modeling and simulation tool for biochemical pathways and cellular processes, and has been used to simulate the activation of pro-subtilisin, the bacterial prototype of pro-furin, as described earlier [66]. The model used in the current simulations is represented in Fig. 4A. Each intermediate has been represented as a discrete molecular species, and reaction pathways defined as reversible or irreversible, with the rate constants noted based on experimental results either with furin, or its homologue subtilisin [66]. Additionally, the software allows for variation of initial concentrations and simulation duration. The models were simulated using the Gillespie stochastic algorithm, with 1000 experiments, $1x10^9$ iterations, for a total duration of 1000 time units. The appearance and disappearance of the various molecular species was monitored and data analyzed in GraphPad Prism.

Results

The constitutively deprotonated mimic of the pH-sensor, H_{69}L-PRO^{FUR}, is more stable than the WT-PRO^{FUR}

To understand the mechanism by which His_{69} functions as a pH-sensor, we undertook detailed structural analyses WT-PRO^{FUR} and the previously reported H_{69}L-PRO^{FUR} variant, which mimics the nonprotonated state of the pH-sensor, using circular dichroism (CD) and intrinsic fluorescence spectroscopy. Prior studies indicate the H_{69}L-PRO^{FUR} chaperones efficient folding of the catalytic domain of furin (MAT^{FUR}), as measured by autoprocessing of the H_{69}L-PRO^{FUR} to form a stable H_{69}L-PRO^{FUR}:MAT^{FUR} complex [103]. However, unlike the WT-PRO^{FUR}:MAT^{FUR} complex, H_{69}L-PRO^{FUR}:MAT^{FUR} remains trapped in a stable state, unable to become active at acidic pH. On the other hand, the H_{69}K-PRO^{FUR} variant, which represents a constitutively protonated state of the pH-sensor, fails to fold correctly and gets rapidly degraded inside the cell [103], and hence was not used in our analyses.

The far UV CD spectrum of the isolated WT-PRO^{FUR} revealed the existence of significant secondary structure (Fig. 3.1A). Substituting the pH-sensor, His_{69}, with Leu caused slight increase in the secondary structure, as seen by the shift in the peak from 206 to 208 nm with a concomitant increase in negative ellipticity at 222 nm in H_{69}L-PRO^{FUR}. As α-helices absorb strongly at 222 nm and 208 nm [156], our results suggest that substituting residues that

mimic the deprotonated state of the pH-sensor marginally increase the α-helicity within the isolated H_{69}L-PRO^{FUR}.

We likewise examined the tertiary structure of the protein by exciting the protein using a wavelength of 295 nm, where the tryptophan emission spectrum is dominant over the weaker tyrosine and phenylalanine fluorescence [168-170]. As seen in Fig. 3.1B, the WT-PRO^{FUR} displays a maximum peak at 342 nm. Under identical conditions, the intrinsic tryptophan fluorescence is slightly enhanced with a blue shift in its emission spectrum (max. at 339 nm) when the His_{69} is substituted by leucine. This indicates that the tryptophan residues are less exposed to solvent when the His_{69} is replaced by a Leu, suggesting the structure may be more packed.

To better understand the extent of stabilization, we next measured the thermodynamic stability of the WT-PRO^{FUR} and its variant, relative to their unfolded states. Thermodynamic stability occurs when a system is in its lowest energy state when compared with all other accessible states within the same reaction environment; it can be measured by monitoring changes in secondary structure with progressive addition of chaotropes such as urea or guanidine hydrochloride [171]. Fig. 3.1C compares chaotrope-induced conformational changes in WT-PRO^{FUR} or H_{69}L-PRO^{FUR} using circular dichroism spectroscopy. The transitions were fitted using a standard Marquardt algorithm with constraints for the baseline set from using the circular dichroism ellipticity of the folded and unfolded proteins [65, 66]. The data demonstrate that H_{69}L-PRO^{FUR} (ΔG_{NU} =1.424 ± 0.12 kcal) is more stable than WT-PRO^{FUR} (ΔG_{NU} =0.921 ± 0.09 kcal/mol). This

indicates the constitutively deprotonated variant H_{69}L-PRO^{FUR} is stabilized by approximately 0.5 kcal/mol when compared to WT-PRO^{FUR}.

Since propeptides are bonafide temporary inhibitors of proteases [5], we next asked how an increase in thermodynamic stability affects the inhibitory function of the isolated WT-PRO^{FUR} and H_{69}L-PRO^{FUR} by comparing IC_{50} values, as described in the Experimental Methods (Fig. 3.1D). Analysis of the data gives an estimated IC_{50} concentration for WT-PRO^{FUR} at pH 6.5 of ~33 nM, three-fold higher than that estimated for H_{69}L-PRO^{FUR} (IC_{50} ~11 nM). This establishes a link between the increased thermodynamic stability of the His_{69}Leu substitution and its ability to act as an inhibitor of MAT^{FUR} as indicated by the decrease in the IC_{50}. Taken together, the circular dichroism and fluorescence spectra, along with the analyses of thermodynamic stabilities suggest that the nonprotonated mimic of the pH-sensor subtly increases both secondary and tertiary structure, and enhances the overall thermodynamic stability and inhibitory function of H_{69}L-PRO^{FUR}.

The co-solvent glycerol enhances structure of the WT-PRO^{FUR} and H_{69}L-PRO^{FUR} and its apparent affinity for MAT^{FUR}

We next examined whether increasing the secondary structure and thermodynamic stability of isolated WT-PRO^{FUR} and H_{69}L-PRO^{FUR} enhances their binding affinity for MAT^{FUR}. We had previously demonstrated that propeptides of aqualysin (PRO^{AQU}) and subtilisins E (PRO^{SUB}), which are orthologs of furin, also exist in partially folded, molten globule-like states [92, 138]. However,

progressive addition of glycerol induced and stabilized the secondary and tertiary structure within PRO^{AQU} and PRO^{SUB}, and simultaneously increased their binding affinities for their cognate catalytic domains [66]. Thus we investigated whether glycerol could stabilize the secondary structures of WT-PRO^{FUR} and H_{69}L-PRO^{FUR}. Our results demonstrate two noteworthy features that are evident in Fig. 3.2A & B; first is the presence of an isosbestic point at a wavelength of 208 nm in WT-PRO^{FUR} and H_{69}L-PRO^{FUR}. The presence of an isosbestic point [172] in the circular dichroism spectra suggests that glycerol induces the partially structured propeptide [140, 173] to fold into a more stable state, and thus there are two distinct states in which the propeptide can exist, depending on its local environment [168]. The second feature to note is the progressive stabilization of secondary structure with increasing amounts of glycerol within isolated WT-PRO^{FUR} and H_{69}L-PRO^{FUR}, measured using changes in ellipticity at 222 nm (Fig. 2C). As previously noted, the secondary structure of H_{69}L-PRO^{FUR} is marginally more stable than that of WT-PRO^{FUR}.

We next examined whether the increase in secondary structure induced by glycerol translates into a tighter binding affinity between the propeptide and the catalytic domain. As seen in Fig. 3.2D, the addition of 30% glycerol enhanced the binding affinity, as measured using IC_{50} values; the IC_{50} of WT-PRO^{FUR} decreased from 33nM to ~2 nM and that of H_{69}L-PRO^{FUR} from 11 nM to ~0.8 nM. Taken together, our results support the hypothesis that both the increased structural (Fig. 3.2A-C) and thermodynamic stability (Fig. 3.1C) of the propeptides enhances their affinity for their cognate catalytic domains.

H_{69}L-PRO^{FUR} is more stable towards pH-dependent unfolding

Since pro-furin undergoes its primary cleavage in the neutral environment of the ER to form a cleaved, non-covalently associated PRO^{FUR}: MAT^{FUR} complex that transits in to the mildly acidic TGN to become active, we next examined how changes in pH affects the structure, stability and binding affinity of WT-PRO^{FUR} and H_{69}L-PRO^{FUR}. Propeptides were purified, refolded, and analyzed for their secondary structure content using CD spectroscopy as described (Experimental Procedures). The results show that WT-PRO^{FUR} undergoes pH dependent unfolding, with an isosbestic point at ~208 nm (Fig. 3.3A). A plot of the changes in CD signal at 222nm as a function of pH suggests that WT-PRO^{FUR} undergoes a cooperative sigmoidal transition to a more unstructured state. Interestingly, the mid-point of this transition occurs at pH ~6.0, close to the optimal pH for activation of furin. Under identical conditions, the H_{69}L-PRO^{FUR}, is more stable; while it does undergo some pH dependent unfolding, with a midpoint of transition likewise at pH ~6.0, it is critical to note that not only it is more stable at neutral pH than the WT, but also that it is not unfolded to the same extent, suggesting that the unfolding response to pH is blunted. In comparison, the change in structure of isolated MAT^{FUR} across this pH range is not significant (Fig. 3.3C, colored lines), nor does the addition of glycerol markedly change the structure of MAT^{FUR} (Fig. 3.3C, grey line). Taken together, this suggests that the protonation status of His_{69} may drive pH dependent conformational changes in the isolated furin propeptide. Our results

indicate that lowering the pH triggers a transition between a folded state at pH 7.4 and a less folded, but not completely unstructured, state at pH 5.0. Mutations of titratable group His_{69} to leucine marginally increase secondary structure at pH 7.4, and to a larger extent at pH 5.0 (Fig. 3.3B), suggesting that protonation of His_{69} is essential for the pH-dependent transition between the two states.

Next we measured the changes in thermodynamic stability of the WT-PRO^{FUR} and H_{69}L-PRO^{FUR}, as described earlier (Fig. 3.1C), under conditions of varied pH. Our data suggest that the overall thermodynamic stability of the proteins decreases when the pH becomes more acidic (Fig. 3.3D). The greater change in thermodynamic stability of H_{69}L-PRO^{FUR} as a function of pH suggested that the H_{69}L substitution enhanced the thermodynamic stability in the isolated propeptide when compared with WT-PRO^{FUR}.

Since the concentration of protons affects the conformation of the WT-PRO^{FUR} and H_{69}L-PRO^{FUR}, we next measured how this conformational change affects the IC_{50} values as a function of pH (Fig. 3.3D). The data demonstrate that the IC_{50} values for WT-PRO^{FUR} and H_{69}L-PRO^{FUR} change as a function of pH, with the maximum inhibitory concentration required for both proteins at pH 6.0 (Fig. 3.3D), the midpoint of the conformational transition ascertained using CD spectroscopy (Fig. 3B). Moreover, three important features in Fig. 3.3D are noteworthy; (i) the IC_{50} value for WT-PRO^{FUR} at pH 7.4 (~12 nM) is about four-fold higher than at pH 6.0 (IC_{50} ~ 50nM); (ii) the IC_{50} value for H_{69}L-PRO^{FUR} at pH 6.0 (~17 nM) is about three-fold lower than that for WT-PRO^{FUR} (IC_{50} ~50 nM); and (iii) when the pH is lower than the optimum for activation (pH~6.0), the IC_{50}

value drops to lower concentrations of propeptides for both WT- PRO^{FUR} and $H^{69}L$-PRO^{FUR}, suggesting an apparent increase in binding affinity. To examine whether pH denatures or inactivates MAT^{FUR}, we also monitored changes in secondary structure (Fig. 3.3C) and activity of furin across this pH range with no propeptide present (Fig. 3.3F). It is worth noting that although the activity of furin does decrease as pH drops, it remains active, with an activity at pH 5.0 roughly 50% of that observed at pH 6.0, where IC_{50} is highest. This suggests that furin remains structurally stable indicating that increased affinity at pH 5.0 is likely a chemical phenomenon. Hence the change in IC_{50}, which is roughly 40-fold lower at pH 5.0 than at pH 6.0, cannot be explained by changes in activity alone. We are currently unable to examine the how pH affects the propeptide: furin complex directly due to the high concentrations of mature furin required to create stoichiometric complexes, we used mathematical simulations to test whether small changes in binding affinities could account for in the inability of the variant to undergo activation, as described in the next section.

Mathematical modeling of furin activation

To better understand the mechanism of furin activation, we modeled the activation of pro–furin complex using CellWare, in a manner similar to our previous work on the activation of pro-subtilisin [66]. The CellWare package offers a multi-algorithmic environment for modeling and simulation of kinetic networks using both deterministic and stochastic algorithms. The software allows modeling of elementary molecular interactions in terms of rate equations, and the temporal changes in molecular species or their stationary state values can be

studied using stochastic and deterministic algorithms. Such models enable the analyses of evolution of molecular species, and how they can affect the overall behavior of the system. The main advantage offered by such models is that it allows us to evaluate how modulating individual interactions affects the overall process outcome and was used to substantiate the experimentally observed stochastic activation of pro-subtilisin [66]. Using CellWare v2.0 we modeled the activation pathway furin. The autoprocessed PRO^{FUR}:MAT^{FUR} complex, free folded PRO^{FUR}, PRO^{FUR}:MAT^{FUR} degradation complex, degraded PRO^{FUR} and free active MAT^{FUR} were represented as individual molecular species. Interactions between them were defined using the law of mass action, and rate constants defined from experimentally determined rates that have been published previously, so as to recapitulate the established timing of furin activation [66, 94]. Fig. 4A depicts the complete pathway. Starting with 10,000 autoprocessed PRO^{FUR}:MAT^{FUR} complexes, the evolution of individual molecular species along the maturation pathway was simulated using the Gillespie algorithm (Experimental Methods). Although mature furin already exists in the cellular environment, our first simulation assumed that furin initially exists only as part of the inhibition complex inside the TGN. In this reaction, the complex has to dissociate into isolated furin and PRO^{FUR} molecules, which can then re-associate in a different configuration to form the Degradation Complex-I (Fig. 3.4A, Reaction 1 and 2, respectively). Numerical values for rates of dissociation of the complex and the formation of Degradation Complex-I were approximated using affinity constants of PRO^{FUR} [14 nM; [103, 174]] and Km for furin substrate

reactions (1.99 μM; [94, 103]). Using this model, the time-dependent activation of mature furin follows a stochastic model and concurs with experimental data for subtilisin, a bacterial homolog of furin [66]. The rapid decrease in accumulated autoprocessed particles coincided with a none-to-all increase in mature furin and is indicative of rapid, autocatalytic activation (Fig. 3.4B). Although the release of free protease is extremely random, a graph of number of active furin molecules versus time of activation for 1000 sample iterations is broadly distributed (Fig. 3.4C, panel 1), with the maximum number of molecules being activated at ~120 min. Thus, by fitting a minimalist mathematical model to the protease activation pathway of furin using a stochastic algorithm, our results appear consistent with our earlier experimental data that suggest the time of activation of furin is approximately 120 minutes [89, 94] However, if active furin already exists when the inhibition complex enters the TGN, which is the likely scenario inside a cell, the distribution of furin molecules versus time is altered, occurring almost instantaneously when even one free molecule is present (Fig. 3.4C panel 2-4). As in this scenario, should the cleavage loop be accessible for cleavage in *trans,* the breakdown of the inhibition complex is rapid and establishes an activation paradigm that is dependent on initial precursor concentration, a case that does not match experimental observations, indicating that there must be some conformational change that modulates access to the cleavage site before processing and activation can occur. We next simulated the activation pathway for different values of K_a (the affinity between the propeptide and protease domain; Fig. 3.4A, Reaction 1) and analyzed the distribution in time of activation

(Fig. 3.4D). For each value of K_a, the mean time of activation for 1000 iterations was estimated. At low K_a, activation is fairly uniform and rapid, whereas higher affinity results in the simultaneous increase in both the time of activation and the associated stochastics (Fig. 3.4D), similar to that observed in the activation of isolated pro-subtilisin [66]. Furthermore, the range of experimentally determined affinities depicted in Fig. 3.4D suggest that approximately 3-fold change in apparent affinity from ~50 nM (WT-PRO^{FUR}) to ~17 nM (H_{69}L-PRO^{FUR}) appears insufficient to account for the experimentally observed inability of H_{69}L-PRO^{FUR}:MAT^{FUR} to undergo activation at acidic pH.. Taken together, our simulations suggest that changes in affinity of the magnitudes seen in our results are not sufficient to account for the inability of H_{69}L-PRO^{FUR} to undergo activation in the secretory pathway.

Another noteworthy observation of our experimental results was the fact that the apparent affinity of the propeptide for the protease domain increases below the optimal pH for activation, despite the decrease in thermodynamic stability of the isolated propeptide. An earlier report suggested that at a pH of 5.0, the catalytic efficiency (k_{cat}) of furin is reduced to 35% of its maximum at optimal pH [175, 176], which may explain the apparent higher affinity seen in our results. To explore this, we again used our model to monitor the time of activation for a complex when the rate of degradation of the IMC (Degradation Complex I and II, Fig. 3.4A, Reaction 4 and 5) is altered. We simulated up to a 10-fold increase and decrease in rate of processing of the IMC, as a surrogate for protease activity; our results show that changes in activity have little effect on the overall

time of activation (data not shown), indicating that a change in activity is unlikely to account for the differences in affinity observed in our studies.

Molecular Dynamics of furin propeptide

Molecular dynamics (MD) simulations can provide information that complements biophysical and biochemical studies of propeptide-mediated protease activation in eukaryotes. MD simulations have been successfully used to model time-dependent changes on residue resolved scale, and the results obtained are consistent with experimental data [92, 148-152, 158, 177]. To investigate how the His_{69}Leu substitution may affect pH-mediated structural changes we employed MD simulations.

We have previously shown that protonation of histidine side chains in the PRO of furin lead to drastic loss of structure during a 10 ns MD simulation [92]. We hypothesized that introduction of the $H_{69}L$ into our model should stabilize the structure. To test this hypothesis we first compared the root mean square fluctuations (RMSF) values of WT-PRO^{FUR} at pH 7.0 and pH 6.0 (Fig. 3.5A). While histidine protonation increased fluctuations at almost all residues, the highest increase was observed in the loop flanked by β2 and β3 in proximity of residue 61, and within the C-terminal half of the cleavage-loop. The largest increase was observed for His_{80}, while His_{69} did not show any change upon protonation and remained stable. The His_{69}leu substitution reduced conformational fluctuations during the simulation at pH 7, with even greater stabilization observed at pH 6. Interestingly, in $H_{69}L$-PRO^{FUR}, His_{80} appears to be the most stabilized compared to residues at pH 6.

Analysis of the root mean square deviation (RMSD) values compared to the starting structure (Fig. 3.5B) as well as ribbon representation of the starting and end structures (Fig. 3.5C), revealed that during the simulation, the core region remains largely stable at pH 7.0 in the WT and H_{69}L-PRO^{FUR}. It is important to note that the loop regions, which had a very high RMSD during the simulation of the WT-PRO^{FUR}, were substantially stabilized by the H_{69}L variant. At pH 6, the core domain of the WT lost its native structure, indicated by rising RMSD values. Compared to WT-PRO^{FUR}, the H_{69}L variant stabilized the core and loop region significantly, although a slight increase in RMSD was still observed.

Hence our MD simulations suggest that while the loop region shows a high degree of flexibility during simulations using both, protonated and unprotonated histidines in WT-PRO^{FUR}, the core remains stable in simulations using unprotonated histidines, but loses structure in simulations using protonated histidines. Introduction of the H_{69}L substitution into our model greatly increased stability of the core region during our simulation using protonated histidines, confirming that protonation of His_{69} alone plays an important role in the pH-mediated structural changes. Interestingly, H_{69}L mutations also lead to increase in stability in the loop regions, which stayed buried during the simulation. Taken together, the MD simulations are consistent with our hypothesis that changes in physical properties of the side-chain at the pH-sensor position has strong influence on the structure of the activation loop and also suggests that leucine, due to its greater hydrophobicity, may not be a faithful representation of unprotonated histidine.

Discussion

In this manuscript, we have employed biophysical, biochemical and computational approaches to investigate pH dependent activation of furin, the canonical proprotein convertase [88]. Our results provide insight into the way in which mature furin recognizes and responds to the changing pH of the secretory pathway, and allow us to propose a mechanism for regulated activation. We also demonstrated that the propeptide provides furin a defined 'activation window', wherein the steps of regulated proteolysis occur within a specific pH range, outside of which the chemical and structural properties of the propeptide appear to block activation of the protease, thereby preventing activation within downstream compartments.

Mechanism of pH dependent activation of pro-furin

There are at least three possible mechanisms (Fig. 3.6) through which the stoichiometric inhibition complex (PRO^{FUR}:MAT^{FUR}) can become active upon reaching the TGN. The first mechanism posits that protonation of the pH-sensor, along with other histidine residues, can induce dissociation of PRO^{FUR} from MAT^{FUR}, which then triggers the second proteolytic cleavage [90, 91]. The second mechanism postulates that pH causes partial unfolding of the propeptide, which enables the second cleavage site to access the active site, promoting proteolysis in a cis-reaction. The third mechanism hypothesizes that protonation of the pH-sensor induces conformational changes that allows a mature furin molecule to access the loop that harbors the second internal cleavage site in

trans [103, 122]. In these cases, subsequent cleavage facilitates propeptide dissociation (Fig. 3.6A).

To test these possibilities, we undertook various *in vitro* and *in vivo* experiments. Our results are most consistent with the hypothesis that the cleavage loop is critical to the overall structure and stability of the propeptide for the following reasons:

(i) The data demonstrate that a 25-fold increase in proton concentration observed between pH 7.5 to 6.0 causes a ~3.5 fold change in apparent binding affinity for WT-PRO^{FUR}. Moreover, H_{69}L pH sensor variant, affects apparent binding affinity ~3.5 fold. To further test this, we built a mathematical model that assumed protonation of the pH-sensor promoted dissociation, allowing it to be processed. By varying K_a, we determined the effect of varying the affinity between propeptide and protease on the rate of activation, which is a stochastic process in bacterial subtilisin [66]. Our simulation results demonstrate that ~10-fold changes in the affinity between the PRO^{FUR} and MAT^{FUR} have only a minimal effect on the rate of activation within physiologically relevant range of values. It is noteworthy that our experimental data establishes the change in affinity due to the substitution is ~three-fold, which argues that dissociation alone cannot account for lack of activation of the H_{69}L-PRO^{FUR}:MAT^{FUR} complex observed in cell-based studies.

(ii) MD simulations suggest that His_{69} protonation affects furin activation by increasing the conformational dynamics of the cleavage loop. At acidic pH, histidine residues within WT-PRO^{FUR} are protonated, including the pH-sensor,

His_{69}. However, when His_{69} is replaced by a Leu, the conformational dynamics of the loop are dramatically reduced, despite all of the remaining histidine residues being protonated (Fig. 3.5). Hence, the protonation status of His_{69} alone affects the dynamics of the activation loop of the wild type and mutant propeptides at two different pHs. Our simulations suggest that at pH 6.0, where the imidazole side chain of histidine is protonated, a dramatic movement in the loop region of PRO^{FUR} precedes the overall unfolding of the propeptide domain. This movement is substantially reduced in case of the $H_{69}L$ variant, resulting in diminished unfolding of the propeptide as seen in Fig. 3.5.

(iii) The addition of the co-solvent glycerol induces greater structure in WT-PRO^{FUR} as compared with compared to the $H_{69}L$-PRO^{FUR}, indicating that the mutant is more ‘native-like’, perhaps because of the hydrophobic packing of the loop into the core of the propeptide. This increased structural stability correlates with an increase in affinity of the propeptide for the protease domain, as evidenced by the lower IC_{50} values calculated from experiments both with the constitutively deprotonated mutant propeptide and with the wild-type as pH is lowered. It is important to note that His_{69} is solvent accessible yet abuts a pocket formed partly by the hydrophobic core residues. Above the pH optima of activation (pH >7.0), the packing of the deprotonated H_{69} into the core maintains a well-packed structure that favors the bound state. However, upon protonation, the imidazole ring becomes charged, disrupting the packing and resulting in destabilization and local unfolding that exposes the cleavage site.

Hence, the results of our experimental and simulated data indicate that structural changes alter the accessibility of the cleavage site, thus raising the question of how the cleavage site becomes available to the active site. While we cannot definitely distinguish between the possibility that the loop movement simply moves the cleavage site into a position more accessible to the active site in *cis*, or that there is a larger destabilization of the packing of the hydrophobic core that allows processing by a second molecule of furin in *trans* (Fig. 3.6), previously reported findings may lend some insight. In earlier work we observed that when the pH-sensor, H_{69}, was mutated to a leucine, no activation of furin takes place under basal conditions [103]. Experiments where excess active furin was added to the inhibition complex indicated that at a non-permissive pH, exogenous furin was unable to affect activation of the inhibited furin. This suggests that the cleavage loop is inaccessible to free furin molecules at a pH outside of its optima [103]. Therefore, we argue that activation is mediated by proteolysis permitted by movement in the cleavage loop that only occurs upon protonation of the pH-sensor, and that dissociation occurs subsequent to processing.

Given this model, it is interesting to consider the possibility that activation is not concomitant with processing; rather, the C-terminal part of the propeptide that sits in the substrate binding pocket, and likely remains bound there for a period of time before it too dissociates to release inhibition. This is consistent with studies that demonstrate the C-terminal propeptides fragments are potent inhibitors of furin [174, 178]. We do not know whether upon cleavage there is a

change in affinity or structure, or if another protease plays a role in the dissociation [179]. We can speculate that after the propeptide is cleaved at R_{75}, the shorter peptide fragment that lies in the substrate binding pocket is simply too short to make efficient contacts with residues of the protease and dissociates, or alternatively, that the cleavage allows a structural change to take place that promotes dissociation. A final alternative possibility is that the peptide fragment then becomes a substrate for cleavage *in trans* by another protease, such as carboxypeptidase [180]. While we cannot yet distinguish between these possibilities, future work will undoubtedly shed further light on this step of activation. Together, the mathematical modeling MD simulations, along with experimentally measured changes in secondary structure and binding affinities are consistent with our model that proteolytic processing precedes propeptide dissociation (Fig. 3.6).

The physical and chemical properties of the propeptide provide an optimal window for pH-dependent furin activation

Another interesting facet of furin activation is the apparent optimal pH-window; just as furin is not activated until it reaches the optimal pH environment of the TGN, it is not active at a lower pH unless it has appropriately transitioned through its activation window. We demonstrate that the propeptide defines this activation window. As indicated by the seeming discordance between pH-dependent changes in thermodynamic stability and affinity for the protease

domain, there are multiple factors at play that determine the inhibitory behavior of the propeptide:

(i) While the thermodynamic stability analyses were conducted on isolated propeptides, the inhibition experiments require association between the propeptide and protease domains, an interaction that we cannot at the current time study directly. Although the isolated propeptide may lose stability at acidic pH, its association with the protease domain may be enhanced due to changes in protonation states of charges side chains in the complex. As depicted in Fig. 7, there are several potential ionic interactions that may affect the interface between the propeptide and the protease domain. Also, given the local hydrophobicity at the interface, the pKa values of the charged groups are likely to be perturbed.

(ii) An alternative possibility for the apparent higher affinity at low pH could be a result of the apparent reduction in enzyme activity due to lower catalytic efficiency as seen in Figure 3.3F and as reported by others [175, 176]. It is noteworthy that although furin activity drops ~50% at pH 5.0 when compared to pH 6.0, the corresponding changes in IC_{50} are ~40-fold over the same pH range. Hence we posit that furin remains structurally stable (as seen by activity and CD studies) indicating that this is likely a chemical phenomenon. We tested this hypothesis further using mathematical simulations, by changing the catalytic efficiency of the enzyme and examining its effects on the time of activation. Our results suggest that in physiologic ranges of the binding affinity, 10-fold changes in the catalytic efficiencies do not significantly alter the time of activation (data not shown). Hence, we argue that the observed optimal window for activation is a

result of the apparent increase in affinity between the propeptide and protease domain. While this may be experimentally tested using isothermal titration calorimetry, the high concentrations of active furin required make these experiments difficult at the present time.

Hence on the basis of our biophysical, biochemical experiments and computational simulations, we argue that the right balance of protonation and destabilization of PRO^{FUR} must be struck in order for efficient proteolysis to occur within the secondary cleavage site within the propeptide. We propose that this balance is only found within the strict confines of the activation window, thus preventing inappropriate activation of the protease within downstream compartments.

Implications of the pH sensor in the activation of proprotein convertases

Data presented here suggest an overarching model for activation of the PCs, using furin, as an example. Upon entering its window of activation, H_{69}, the pH-sensor in furin gets protonated, to destabilize the hydrophobic pocket in which the pH-sensor sits. Destabilization pushes the cleavage loop outward, thus allowing the catalytic site access to the secondary cleavage site. At the same time, a certain amount of flexibility in the association of the propeptide with the protease domain is preserved, which allows this cleavage to act as a finely tuned trigger. Above the pH optima for activation, the stability of the propeptide due to the hydrophobic packing of the pocket keeps the propeptide tightly associated with the protease, and it is likely that below this window, secondary associations

form as acidic residues, such as glutamates and aspartates are protonated, thus forming salt bridges that keep the propeptide likewise tightly associated. Hence, it is only within a small window of pH that furin is able to be activated.

We have demonstrated that the propeptides of furin and PC1 alone contain information necessary for their compartment specific activation [92]. However, the residue that corresponds to His_{69} in furin is conserved within all PCs, suggesting that additional factors must augment the subtle differences between the pH optima of individual PCs. This may be in part mediated by the distribution of additional histidines and other charged residues within the propeptides of PCs. Whether such evolutionary enrichment in His content is evident in other protease families is a subject of ongoing research in our laboratory.

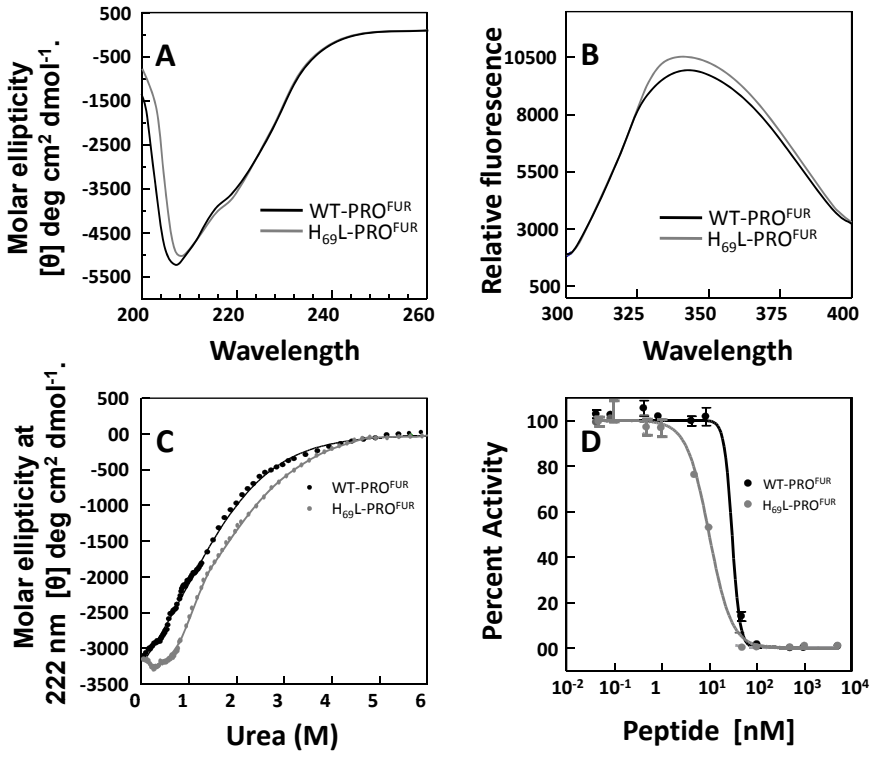


Figure 3.1: $H_{69}L\text{-}PRO^{FUR}$ is more structured than $WT\text{-}PRO^{FUR}$
(A) Secondary structure of $WT\text{-}PRO^{FUR}$ and $H_{69}L\text{-}PRO^{FUR}$ determined via CD spectroscopy at far UV, performed at pH 7.0, and plotted as molar ellipticity [θ] deg cm^2 $dmol^{-1}$. **(B)** Tertiary structure of wild type or mutant propeptide determined by measuring intrinsic tryptophan fluorescence after excitement with λ=295 nm. **(C)** Thermodynamic stability of the propeptides monitored by changes in ellipticity, [θ], at λ=222 nm as a function of urea concentration. Data were fit to a standard two-state equation using a Marquardt algorithm. **(D)** Normalized activity, used to estimate IC_{50} values, determined by monitoring cleavage of the fluorogenic peptide substrate Abz-RVKRGLA-Tyr[2-NO_2], with increasing amounts of $WT\text{-}PRO^{FUR}$ or $H_{69}L\text{-}PRO^{FUR}$ present. All data are averaged over three independent experiments.

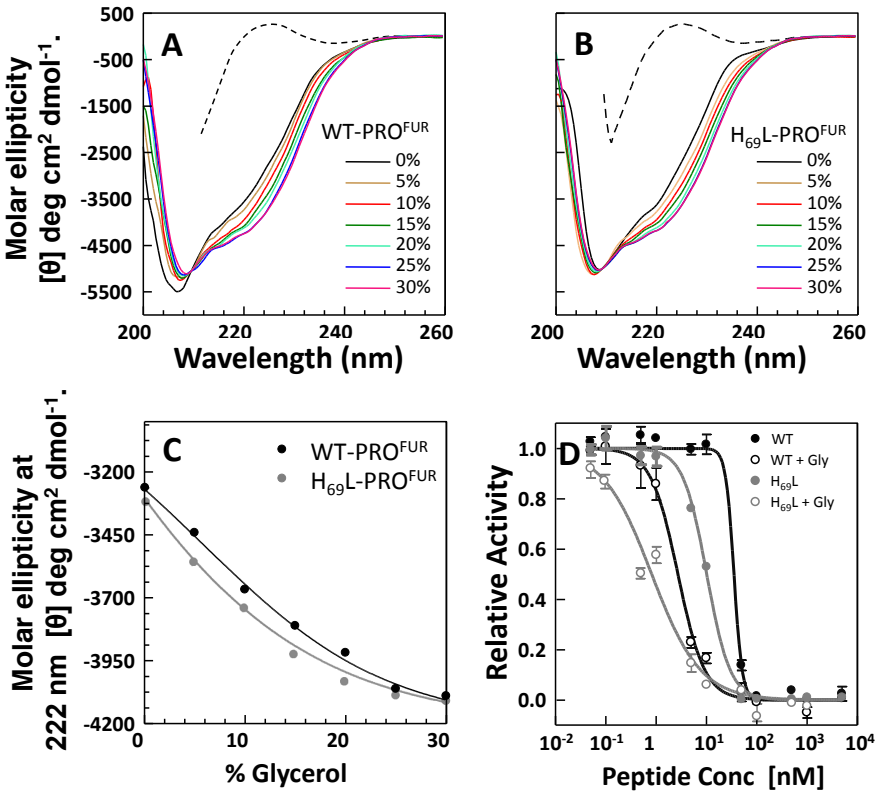


Figure 3.2: Glycerol induced structure of the isolated propeptides enhances affinity for MAT^{FUR}.
Secondary structure of **(A)** WT-PRO^{FUR} and **(B)** H_{69}L-PRO^{FUR} measured using circular dichroism spectroscopy as a function of increasing concentration of the co-solvent glycerol. Data are plotted as molar ellipticity. **(C)** Changes in secondary structure of the isolated propeptides, monitored by changes in ellipticity at λ=222 nm, with increasing concentration of glycerol. **(D)** Activity, normalized to maximal activity with no PRO, used to estimate IC_{50} values of the propeptides in the presence (empty circles) or absence (filled circles) of 30% glycerol.

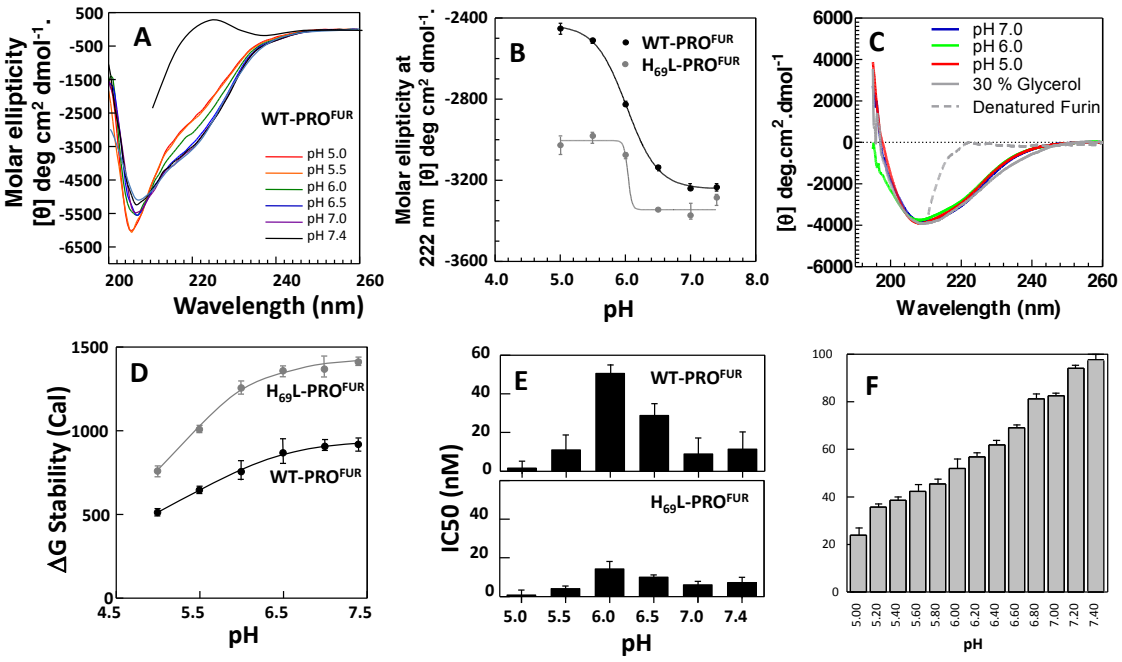


Figure 3.3: H_{69}L-PRO^{FUR} is more stable than WT-PRO^{FUR} to pH-induced unfolding.
(A) pH-dependent secondary structure of WT-PRO^{FUR} performed at pH 7.0-5.0 and plotted as molar ellipticity. **(B)** Changes in secondary structure of the isolated propeptides, monitored by changes in ellipticity at λ=222 nm, plotted as a function of increasing pH. The midpoint of the unfolding transition for both peptides occurs at pH ~6.0. **(C)** CD structure of MAT^{FUR} at varying pH (colored lines) and with the addition of 30% glycerol (grey line). The dotted line represents the spectra of denatured MAT^{FUR} **(D)** Thermodynamic stability of WT-PRO^{FUR} and H_{69}L-PRO^{FUR} as a function of pH. **(E)** IC_{50} values for WT-PRO^{FUR} (top) and H_{69}L-PRO^{FUR} (bottom), as a function of pH. **(F)** Activity of furin in the absence of the propeptide at varying pH; all values given as percentage of maximum activity, and are the average of 3 independent experiments.

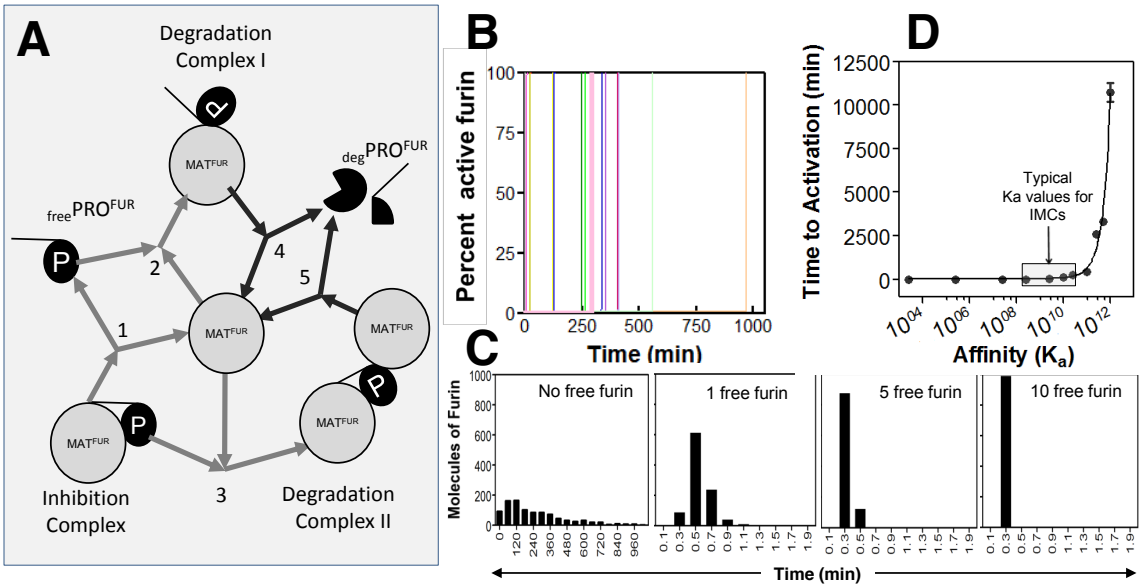


Figure 3.4: Mathematical modeling of furin activation
(A) Model of furin activation pathway used in mathematical simulations. Fur = Furin (protease), P = PRO (propeptide). Rates for each step were modeled in accordance with published values in the literature. The grey and black arrows depict reversible and irreversible reactions, respectively. **(B)** Stochastic activation of furin over 15 sample simulations. **(C)** Distribution of stochastic activation times (min) of furin over 1000 sample iterations (Left-most, No free furin) and distribution of activation times (min) when active furin is present (left-middle, right-middle, and right-most). **(D)** Change in time to activation (min) as K_a, the affinity of Furin for free PRO^{FUR}, is altered. Grey box indicates physiologic range. **(E)** Change in time to activation (min) as k_{cat}, the catalytic efficiency, or the rate of dissociation of deg-PRO^{FUR} from Degradation Complex I and II to release active Furin, is altered.

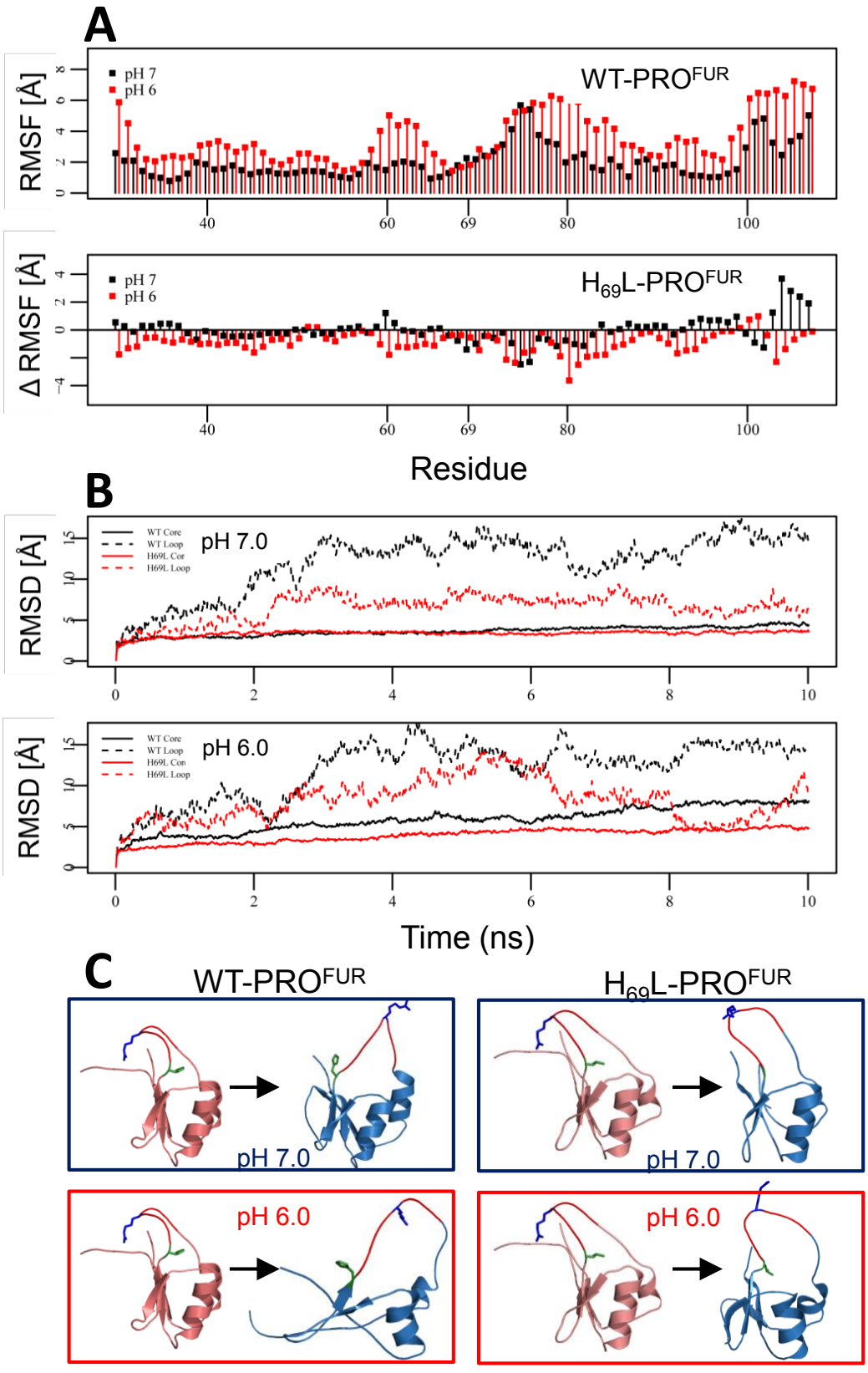

A
RMSF [Å]
pH 7
pH 6
WT-PRO^FUR
40
60
69
80
100
Δ RMSF [Å]
H69L-PRO^FUR
Residue
B
RMSD [Å]
WT Core
WT Loop
H69L Cor
H69L Loop
pH 7.0
pH 6.0
Time (ns)
C
WT-PRO^FUR
H69L-PRO^FUR
pH 7.0
pH 6.0

Figure 3.5: Molecular Dynamics indicates protonation of H69 destabilizes the core of PROFUR

MD simulations were performed on WT-PRO^{FUR} and H_{69}L-PRO^{FUR} using NAMD as described in methods. **(A)** Top: RMSF as a function of residue number for WT-PRO^{FUR}. Bottom: Differences between the RMSF of simulation with H_{69}L-PRO^{FUR} and the WT (ΔRMSF). Negative values indicate reduced fluctuations and positive values indicate increased fluctuations due to the H_{69}L point mutation. Values obtained under simulated pH of 7 are shown in black, while values obtained under simulated pH of 6 are in red. **(B)** RMSF values for core region (all except loop) are shown by dashed lines, and loop regions (residues 70-79), are depicted by solid lines and plotted as a function of simulation time. Black lines represent WT-PRO^{FUR} and red lines H_{69}L-PRO^{FUR}.**(C)** Ribbon representation of the starting (red) and final (blue) structures of the simulations. The secondary cleavage site, R_{75} is indicated in the cleavage loop (blue), and the pH sensor, H_{69}, indicated in green. All simulations were done over 10ns.

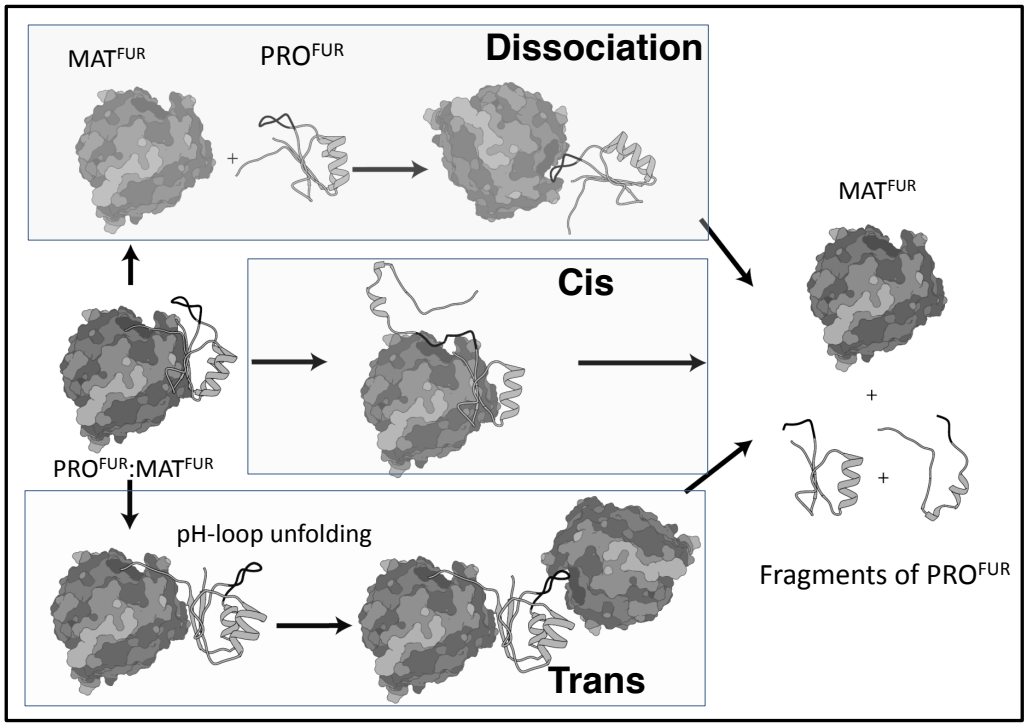


Figure 3.6: Models of Furin Activation
Three potential models of furin activation considered. The dissociation model **(top)** posits that PRO^{FUR} dissociates from MAT^{FUR} after the pH-sensor is protonated, then reassociate in a different orientation such that the propeptide can be cleaved. The cis and trans models of processing **(middle and bottom, respectively)**, in contrast, both suggest that protonation of the pH sensor drives an unfolding event that allows the cleavage site loop to become accessible to the active site, either of its own MAT^{FUR} to be cleaved in cis **(middle)**, or to a second molecule of MAT^{FUR} to be cleaved in trans; only after the propeptide is cleaved does it dissociate from the protease.

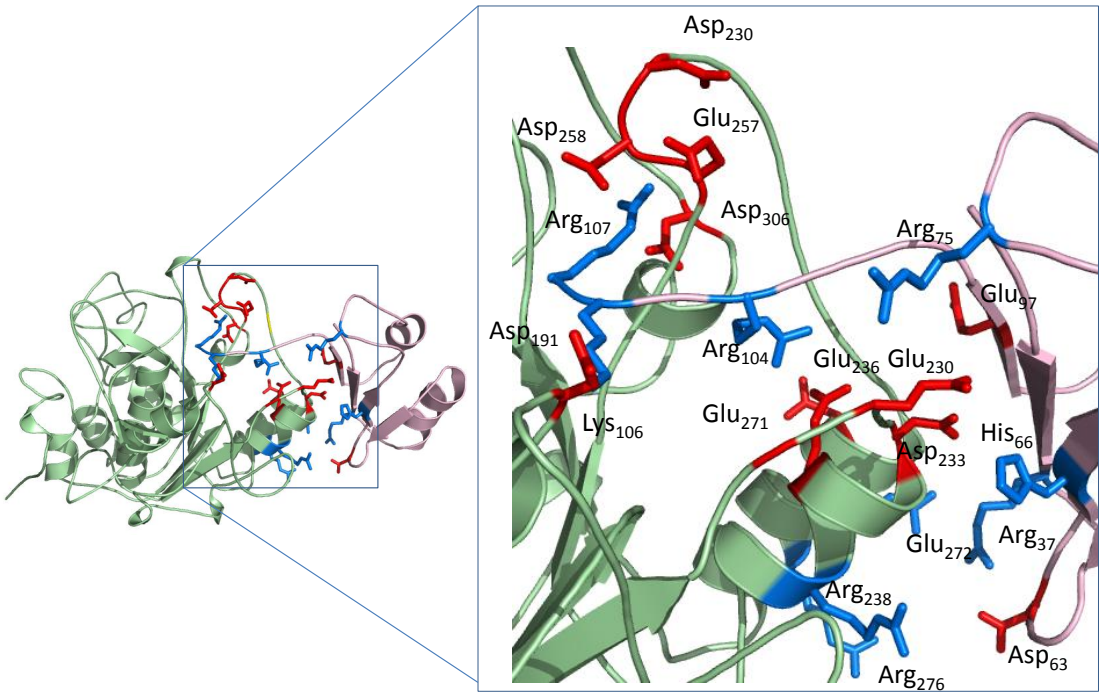


Figure 3.7: Interface between MATFUR and PROFUR
Ribbon diagram showing potential interaction of residues at the interface of MAT^{FUR} (green) and PRO^{FUR} (pink). Acidic residues are indicated in red, and basic residues in blue.

Chapter IV: Mechanism of fine-tuning pH sensors in Proprotein Convertases: Identification of a pH-sensing histidine pair in the propeptide of proprotein convertase 1/3

Danielle M. Williamson, Johannes Elferich, and Ujwal Shinde

From the Department of Biochemistry and Molecular Biology

Oregon Health and Science University, Portland, OR 97239

This chapter was published in The Journal of Biochemistry in July 2015. Experiments and manuscript preparation were done by DMW, with assistance in figure preparation and editing from JE and US.

Abstract

The propeptides of proprotein convertases (PCs) regulate activation of cognate protease domains by sensing pH of their organellar compartments as they transit the secretory pathway. Earlier experimental work identified a conserved histidine-encoded pH sensor within the propeptide of the canonical PC, furin. To date, whether protonation of this conserved histidine is solely responsible for PC activation has remained unclear, due to the observation that various PC paralogues are activated at different organellar pH. To ascertain additional determinants of PC activation, we analyze Proprotein Convertase 1/3 (PC1/3), a paralogues of furin that is activated at a pH of ~5.4. Using biophysical, biochemical and cell-based methods, we mimicked the protonation status of various histidines within the propeptide of PC1/3, and examined how such alterations can modulate pH-dependent protease activation. Our results indicate that while the conserved histidine plays a crucial role in pH sensing and activation of this protease, an additional histidine acts as a “gatekeeper” that fine-tunes the sensitivity of the PC1/3 propeptide to facilitate the release inhibition at higher proton concentrations. Coupled with earlier analyses that highlighted the enrichment of the amino acid histidine within propeptides of secreted eukaryotic proteases, our work elucidates how secreted proteases have evolved to exploit the pH of the secretory pathway by altering the spatial juxtaposition of titratable groups to regulate their activity in a spatiotemporal fashion.

Introduction

Eukaryotic cells have evolved an elegant series of membranous compartments to regulate protein synthesis, folding, activation, sorting and export [136, 137]. Beginning with their entry into the endoplasmic reticulum (ER) as they exit the ribosome, proteins transit these secretory pathway compartments on their journey to their ultimate destination. Just as each group of compartments have mechanisms for maintaining precise pH and calcium balance [181, 182]; each family of proteins has co-evolved means to sense the environmental changes to ensure optimal organismal homeostasis. How eukaryotic proteins have diverged from their prokaryotic ancestors to exploit the unique organellar environment of the secretory pathway for biological function remains a fundamental question in cell biology [122].

Proprotein convertases (PCs) are eukaryotic members of the ubiquitous super-family of subtilases [88]. Comprised of nine serine endoproteases (PC1/3, PC2, furin, PC4, PACE4, PC5/PC6, PC7/LPC/PC8, SKI/S1P, and NARC1/PCSK9), PCs are responsible for the conversion of a diverse range of precursor substrates to their active forms, and thus play a central role in maintenance of physiologic homeostasis within cells and tissues [122, 183]. Furin, the most thoroughly characterized PC, is constitutively expressed in virtually all tissues, and catalyzes the maturation of a diverse repertoire of hormones, enzymes, and receptor precursors within the secretory pathway [88]. Not surprisingly, misregulation of furin results in both hyper- and hypo-activity,

and has been associated with cancer invasiveness and metastasis [184-189], susceptibility to viral and parasitic infection [190], and increased severity of cardiovascular disease [191, 192]. Additionally, furin-deficient mice die at embryonic day 11 due to cardiac defects resulting from failed chorioallantoic fusion, axial rotation and ventral closure [193]. PC1/3 (also known as PC1 or PC3), along with PC2, is a neuroendocrine convertase responsible for the processing of several critical metabolic regulators, including insulin, glucagon, and proopiomelanocortin (POMC) [194]. Knockout mice of *PCSK1* or *PCSK2*, the genes encoding PC1/3 and PC2 respectively, remain viable despite hormonal and/or neuroendocrine deficiencies [126, 162, 195]. Consistent with this, several studies characterizing mutations in PC1/3 have demonstrated an altered substrate processing that may underlie obesity, type II diabetes mellitus, and endocrine derangements [126, 196-199]. Further support for the correlation between polymorphisms in PC1/3 and metabolic disease is offered by reports of both individual patients [199-203] as well as epidemiologic studies [195].

As unregulated protease activity can have devastating consequences on organismal homeostasis [204], PCs, like all subtilases, are synthesized as zymogens, and are activated in a precise spatiotemporal fashion by intramolecular proteolysis [66]. A fundamental question has long been how the timing of activation is encoded and recognized by the protease. Our understanding of how this regulation is achieved is based on studies of profurin [92, 95, 104, 205]. This furin precursor contains an 83-residue N-terminal propeptide that is requisite for folding of its cognate catalytic domain in the ER

[89, 94, 164], after completion of which the propeptide gets cleaved but subsequently remains associated as an inhibitor until reaching the trans-Golgi network (TGN). The necessity of furin to reach the TGN before becoming active [165], coupled with the understanding of the unique environments of each compartment of the secretory pathway [206], argued for the presence of an encoded sensor that recognizes and responds to specific environmental signals [92]. In fact, studies demonstrate that it is the mildly acidic pH of the TGN (pH ~ 6.5) that triggers the release and degradation of the furin propeptide (PRO^{FUR}), and thus release of inhibition of the protease domain of furin (MAT^{FUR}) [89, 94]. Interestingly, while the PC1/3 precursor, proPC1/3, transits the secretory pathway in much the same way as the furin precursor, our previous work indicates that PC1/3 requires a lower pH for its activation (pH ~ 5.5), and thus is likely activated in a later compartment [92].

Organellar pH can alter the protonation status of charged residues, thus altering the structure and stability of the protein both on a local and global scale [207]. Given that the pH range of interest in the case of the secretory pathway and PC activation falls within the physiologic range of ~7.4 in the ER to ~5.4 in DCSGs, altered protonation of basic residues such as arginine (pKa ~12.5) or acidic ones such as glutamate (pKa ~4.2) were less likely candidates for pH sensors; however, the imidazole ring of histidine has a pKa of ~6.0, making it ideally situated to respond to pH changes in this range. Histidine is used as a pH sensor in a variety of biological molecules, including hemoglobin and class II MHC [98, 208, 209], and is enriched within the propeptides of eukaryotic

proteases that transit the secretory pathway, in contrast to both cytosolic proteases and prokaryotic orthologues [95, 205]. Indeed, we identified a histidine (His_{69}) within PRO^{FUR} that regulates the compartment-specific activation of MAT^{FUR} [103]. This histidine is in a loop adjacent to the secondary cleavage site, nestled in a solvent accessible pocket lined by hydrophobic residues. At the near-neutral pH of the ER, the deprotonated histidine acts as a hydrophobic residue, stabilizing the packing within this pocket, and keeping the pH-sensitive loop protected against cleavage. Upon entry into the TGN, the histidine is exposed to a ~10-fold higher proton concentration, and thus is protonated; as a result, the imidazole side-chain becomes polar, disrupting the packing, and driving a local conformational change that exposes the secondary cleavage site, allowing for rapid degradation and release of PRO^{FUR} from its cognate protease domain, MAT^{FUR}[104].

Propeptide domains alone encode sufficient information for regulating the organelle-specific pH-dependent activation of cognate protease domains [92]. Thus, swapping of propeptides between furin and PC1/3 transfers pH-dependent protease activation in a propeptide-dictated manner *in vitro* and in cells. Interestingly, the histidine pH sensor identified in PRO^{FUR} is absolutely conserved within the propeptides of all PCs, including $PRO^{PC1/3}$. Despite their structural similarity and the presence of the conserved histidine throughout the family of PCs, it remains unclear how the pH sensitivity of other PCs is encoded.

In this manuscript, we ask, if the pH-sensing histidine is conserved in $PRO^{PC1/3}$, why does $MAT^{PC1/3}$ not become active at the same pH as furin? We

present biochemical and structural data that suggest the conserved histidine residue plays a critical role in pH dependent activation, similar to what has been established in furin; however an additional histidine residue modulates this pH sensitivity such that a more acidic environment is required for PC1/3 activation.

Materials and methods

Cell culture and transient transfection:

Cos7 cells were maintained in Dulbecco's modified Eagle high glucose medium (Hyclone) containing 10% v/v fetal bovine serum and 1% v/v penicillin-streptomycin. Cells were kept at 37^{o}C in a 5% CO_2 environment, as previously described [92, 104]. For expression of ER retained constructs, Cos7 cells at 60-80% confluence were transfected with pcDNA3.1 expression vectors containing WT PC1/3 truncated at Arg_{618} with a C-terminal –KDEL sequence, [92] or vectors containing the noted $PRO^{PC1/3}$ variants using TransIT LT-1 transfection reagent (Invitrogen) as per manufacturer instructions. The constructs also contained an HA tag at the N-terminus of the propeptide and a FLAG-tag at the N-terminus of the protease domain (Figure 1A).

pH-dependent Activation Assays:

Twenty-four hours post-transfection with ER-retained PC1/3 constructs, cells were washed 2x in PBS and incubated for 15 minutes in fractionation buffer [270 mM sucrose, 10 mM Tris, pH 7.4] with protease inhibitors (cOmplete EDTA free, Roche). Cells were lysed with 3 rounds of sonication (10 x 1 sec pulses at

30% power) in an ice bath. Unlysed cells and cell debris was pelleted at 800x g for 15 minutes at 4°C, and supernatant transferred to a clean tube. Samples were incubated with 30% w/v PEG8000 for 16 hours at 4°C with gentle agitation. Precipitated proteins were pelleted by centrifugation at 17000x g for 30 minutes at 4°C. Pellets were resuspended in activity assay buffer [100mM NaOAc, 0.1% Brij-35, 5mM $CaCl_2$] at a pH of 7.4, 6.4 or 5.4, and incubated at 37°C for 1 hour to allow for activation, unless otherwise noted. After 1 hour, 15 μM of the fluorogenic substrate Abz-RVKRGLA-Tyr[3-NO_2] was added, activity assayed in triplicate on a SpectraMax-M2 spectrofluorometer equipped with a 96-well plate reader (Ex 320/Em 425) for 1 hour at room temperature. Data were fitted and analyzed using GraphPad prism. Replicate samples were also processed for western blot; the anti-FLAG M2 antibody (Sigma) was used to recognize the N-terminally FLAG-tagged mature protease, and anti-HA.11 (Covance) used to recognize the N-terminally HA-tagged propeptide.

For experiments where limited trypsinization was used to digest the propeptide, PEG8000 precipitated proteins were resuspended in activity assay buffer, pH 7.4 as above, and incubated in the presence or absence of agarose-immobilized TPCK-trypsin (~10 TAME units per sample, Thermo Scientific) for 1 hour 37 °C. Samples were then treated with soybean trypsin inhibitor to a final concentration of 1 mg/ml for 20 min at 37°C before proceeding with the activity assay as above.

Protein Production and Purification:

Codon-optimized sequences encoding either WT-PRO$^{PC1/3}$ (mouse) or leucine and arginine variants were cloned into the pET11b backbone and expressed in BL21(DE3) E. coli as previously described [55, 92]. Protein was purified from the soluble fraction by ion exchange after cell lysis via French pressure cell, and dialyzed into 6 M guanidinium HCl-containing buffer [50 mM Tris-HCl, 6 M Guanidinium HCl, pH 6.5] for long term storage. Before use, proteins were refolded by dialysis against refolding buffer [10 mM Tris, 10 mM Cacodylate, 10 mM NaOAc, 150 mM KCl, 5 mM $CaCl_2$] pH 6.8 unless otherwise noted. Concentration was determined after refolding by absorption at 280 nm.

Enzymatically active mature PC1/3 was expressed via the baculovirus system. The coding sequence of mouse PC1/3 was cloned into the pAcGP67A vector (BD Biosciences), using the primer pair TTTGCGGCGGATCCCGGGAAGAGGCAGTTTGTTAATGAATG and TTCCATGCGGCCGCTCATCATCTCCTGTCATTCTGGACTGTATT into the XmaI and NotI sites, resulting in replacement of the endogenous signal sequence with the gp67 signal sequence and truncation at R_{618}. The resulting coding sequence was subcloned into a pFastBac transfer vector (Life Technologies) containing a C-terminal GFP and 8X His tag.

The baculovirus genome was generated by transfection into DH10Bac cells according to manufacturer protocols (Bac-to-Bac, Life Technologies). P1 virus was generated by infection of adherent Sf9 cells and amplified to the P3

as described in the manual. Virus titer was measured by endpoint dilution and quantification of infected cells by GFP fluorescence.

Protein was produced by infecting Sf9 suspension cultures in SF900-III medium (Life technologies) at 4x10^6 cells/ml at a MOI of 2. After 24h, the media was clarified by two subsequent centrifugation steps at 500x and 10,000x g, respectively, for 20 minutes at 4^{o}C. Protein was precipitated using 20% w/v PEG-8000 and resuspended in 100-fold smaller volume of 50 mM MOPS at pH 6.5, 300 mM NaCl, 5 mM $CaCl_2$, and 0.01% v/v Brij-35 (Buffer A) mixed with fresh cOmplete EDTA-free Protease Inhibitor cocktail (Roche Diagnostics). After clarification by centrifugation at 10,000x g for 20 minutes, 10 mM imidazole was added and the buffer was agitated for 2h with 1 ml NiNTA resin (Thermo Scientific). Afterwards the buffer was packed into a column and extensively washed (>20x column volumes) of Buffer A with 25 mM imidazole. PC1/3 was then eluted using a linear imidazole gradient to 500 mM and fractions containing GFP fluorescence and PC1/3 activity were pooled, flash-frozen, and stored at -80^{o}C (Figure 4.1B). The fractions indicated were pooled, with an estimated yield of 100 mg/L and a purity of 90%. A representative enzyme progress curve for PC1/3 is depicted in Figure 1C. While PC1/3 is known to exhibit complex enzyme kinetics, including a lag phase in activation, we see minimal evidence of this at the enzyme concentrations used in our studies.

Circular dichroism spectroscopy:

Circular dichroism spectroscopy was performed on an AVIV model 215 CD spectrometer at 4^{o}C as previously described [92, 104]. Briefly, after

centrifugation at 100,000x g for 30 minutes, protein concentration was adjusted to approximately 0.3 mg/ml and spectra obtained using a 1 mm quartz cuvette for far UV spectra. A minimum of three independent scans were acquired and data averaged [65, 66, 210]. Structural changes due to substitutions at individual histidine residues in $PRO^{PC1/3}$ were monitored by plotting the changes in ellipticity at 222 nm at different pH using the WT $PRO^{PC1/3}$ as a control. Equilibrium unfolding can be represented as contributions of two states, the native (N) and the unfolded (U) as previously described [92, 210], given by:

$$A_{obs} = \frac{A_N + A_U exp^{\frac{-(\Delta G_{NU} - m_{NU}c)}{RT}}}{1 + exp^{\frac{-(G_{NU} - m_{NU}c)}{RT}}}$$

Alternately, the Henderson-Hasselbalch equation was used to describe the unfolding, as originally described by Tanford [211]:

$$A_{obs} = A_U + (A_N - A_U) * \frac{K_a/[H^+]^n}{1 + (K_a/[H^+]^n)}$$

where n = 1 or 2 for a one-proton or two-proton model.

Activity assays and determination of IC_{50}: Activity assays were performed to determine the inhibitory capabilities of the various propeptides as previously described [92, 104]. Briefly, 15 μM of the fluorogenic substrate (Abz-RVKRGLA-Tyr[3-NO_2]) was incubated with serially diluted amounts of either refolded WT or mutant $PRO^{PC1/3}$ in concentrations ranging from 0.15-3000nM) in activity assay buffer [100 mM NaOAc, pH 6.8, 0.1% v/v Brij-35, 5 mM $CaCl_2$], and 0.02 EU [1

EU of enzyme produces 267 RFU/min in the assay buffer] of mature PC1/3 was added to initiate the reaction. Activity was assayed as described above. Data were fitted and analyzed using GraphPad prism to determine IC_{50} values as previously described [92, 104].

Results

Sequence and structural analysis of PRO^{FUR} and $PRO^{PC1/3}$ orthologues: Multiple sequence alignment of PRO^{FUR} and $PRO^{PC1/3}$ provides several insights that serve as a starting point to generate a hypothesis about the mechanism of pH sensing in PC1/3, and how differences in sequence can encode different sensitivities to pH between the two homologues. The alignment of eight orthologues of PRO^{FUR} and $PRO^{PC1/3}$ indicates a high degree of sequence similarity, including stretches of completely conserved residues distributed throughout the entirety of both propeptides (Figure 4.2A). A structural comparison establishes that both propeptides adopt almost identical folds and are comprised of two alpha helices packed against four beta sheets, with a flexible loop region containing the secondary cleavage site between helix 1 and 2 (Figure 4.2B and C). As we have previously established that the pH sensor in furin is a histidine residue located at position 69 in the sequence [103], we paid particular attention to the distribution of histidines within $PRO^{PC1/3}$ compared to PRO^{FUR}. While PRO^{FUR} has histidines distributed throughout, $PRO^{PC1/3}$ has only four, all of which are clustered within 10 residues of the secondary cleavage site. Of these, His_{72} in $PRO^{PC1/3}$ aligns with His_{69}, the pH sensor identified in PRO^{FUR}

and is positioned similarly at the N-terminus of the cleavage loop, thus was of primary interest. Notably however, His_{72} appears more buried within the core of $PRO^{PC1/3}$ than His_{69} in the PRO^{FUR} (Figure 4.2B), and His_{75} in $PRO^{PC1/3}$, which does not have a corresponding histidine in PRO^{FUR}, is also located within the cleavage loop of $PRO^{PC1/3}$. Based on the structure [1KN6, [91]], His_{75} in $PRO^{PC1/3}$ is surface exposed and appears to overlie His_{72}, leading us to speculate that His_{75} may modulate the solvent accessibility of His_{72}. Two other histidines, His_{67} and His_{85} are also located within close proximity of the cleavage site loop; His_{67}, which corresponds to His_{66} in furin, is N-terminal to the cleavage loop, and is situated at the interface that makes contact with the protease domain. Although His_{66} was proposed as a potential pH sensor [90, 91], subsequent biochemical and cell-based experiments demonstrated that this residue does not play a role in the pH-dependent activation of furin [103], making the corresponding His_{67} in $PRO^{PC1/3}$ a less likely player in the pH sensing mechanism of this protease. The remaining histidine residue (His_{85}) does not contribute significantly to the packing of the core of the propeptide and is solvent exposed within helix 2 located at the C-terminal side of the cleavage loop. Since local loop movement is critical for facilitating accessibility of the secondary cleavage site for proteolysis, His_{85} appears to be an unlikely candidate for a pH sensor, a supposition that is consistent with our preliminary IC_{50} values and circular dichroism data that showed the protonation at these positions do not have significant impact on the binding, structure or stability of the isolated $PRO^{PC1/3}$ (data not shown). Hence, in

this manuscript we focus on His_{72} and His_{75} as putative pH sensors within $PRO^{PC1/3}$.

Analysis of the role of histidine residues in the pH dependent activation of the precursor protein, proPC1/3: To examine the role that the two histidine residues (His_{72} and His_{75}) may play in the pH-dependent activation of the precursor proPC1/3 *in vivo*, we first measured the ability of propeptide variants of PC1/3 to undergo activation using ER-localized constructs as previously described [Figure 1A [92]]. Here proPC1/3 (~78kDa) undergoes efficient auto-processing to produce $MAT^{PC1/3}$ (~67kDa) that remains non-covalently associated with $PRO^{PC1/3}$ (~10kDa) to form a stable inhibition complex [92]. Since the KDEL motif restricts proteins to the neutral environs of the ER, $PRO^{PC1/3}$ in the inhibition complex does not undergo the second autoprocessing step that is necessary for degradation and release of the propeptide. Hence, at a pH of 7.4 and 6.4, the ER extracts of $MAT^{PC1/3}$ fail to degrade the $PRO^{PC1/3}$ from the inhibition complex (Figure 3A, inset), and displays negligible catalytic activity against the fluorogenic peptide substrate (Figure 4.3A). However, preincubation of the ER extracts of $MAT^{PC1/3}$ at pH 5.4 causes a robust increase in the catalytic activity (Figure 4.3A), which coincides with degradation of $PRO^{PC1/3}$, as seen by the Western blot analysis (Figure 4.3A inset). It should be noted that this second autoprocessing event is time dependent, as degradation of the propeptide does not take place immediately upon exposure to low pH, and at longer times of incubation, continues to be degraded (Figure 4.3E).

We then compared the WT proPC1/3 with variants in which His_{72} or His_{75} individually, or both His_{72} and His_{75} together were substituted with either a leucine or arginine, to mimic the constitutively deprotonated or protonated states of histidine, respectively, in accordance with our earlier work on the pH sensor in PRO^{FUR}[92, 104]. Substituting either His_{72} or His_{75} with leucine allows for proPC1/3 variants to undergo processing at the primary cleavage site to produce $MAT^{PC1/3}$ (Figure 3D), but abrogates degradation of the histidine to leucine variants of $PRO^{PC1/3}$ at pH of 5.4, thereby preventing the catalytic domain, $MAT^{PC1/3}$ from cleaving the fluorogenic substrate (Figure 4.3A and inset). Furthermore, no additional decrease in activity is observed when both histidines were mutated to leucine.

To confirm whether $MAT^{PC1/3}$ obtained using the His_{72} or His_{75} $PRO^{PC1/3}$ variants are catalytically active, the inhibition complexes were treated with trypsin (Figure 4.3C) to cleave the arginine and lysine rich propeptides. The results establish that the inhibition complexes, which fail to undergo auto-activation at pH 7.4, can be activated by incubation with trypsin at this pH, indicating that the protease itself is functional and that the His_{72} or His_{75} $PRO^{PC1/3}$ variants are incapable of undergoing pH-dependent activation.

In contrast with the leucine substitutions, when His_{72} or His_{75} were replaced with arginine, the activity of the resulting $MAT^{PC1/3}$ was more nuanced (Figure 4.3B). While the double mutant, $His_{72/75}$Arg appears constitutively active at all pHs, the maximal activity of the His_{72}Arg variant is only marginally greater than the control at all pHs, suggesting that protonation at this position, while

necessary (as evidenced by the inability of the His_{72}Leu variant to be activated) must be appropriately timed. The western blot depicting the processing of the precursor (Figure 4.3D), demonstrates that the His_{72}Arg proPC1/3 variant does not undergo appreciable processing and suggests that inappropriate protonation of this residue may impair folding and/or processing, similar to results seen in furin [103]. An alternate explanation is that this mutation in the propeptide impairs autocatalytic activity, resulting in reduced propeptide excision in the ER, similar to the previously reported Ser_{307}Leu mutation, which also displayed diminished activity *in trans* against a peptide substrate [199]. Interestingly, the His_{75}Arg variant shows activity at pH 7.4 that is higher than that of the wild type, and appears to be completely active at pH 6.4, making the resulting $MAT^{PC1/3}$ more furin-like. Not surprisingly, the western blot analysis establishes that the His_{75}Arg variant of proPC1/3 is efficiently autoprocessed and $H_{75}R\text{-}PRO^{PC1/3}$ is completely destroyed at pH of 7.4 on a timescale that we were not able to capture in our analysis (Figure 4.3B and inset). Thus, our mutational analysis suggest that both histidines at position 72 and 75 within $PRO^{PC1/3}$ may be playing roles in the pH-dependent activation of proPC1/3 as it transits the secretory pathway.

Effect of pH on structural changes in $PRO^{PC1/3}$: Since the protonation status of the pH sensor alters the structure of PRO^{FUR} [92, 104], we next analyzed how the protonation of His_{72} and His_{75} affects the secondary structure of isolated $PRO^{PC1/3}$ by circular dichroism spectroscopy using leucine and arginine variants as mimics for the constitutively protonated and deprotonated states, respectively. The CD spectra reveal the presence of significant secondary

structure within isolated $PRO^{PC1/3}$, which is not significantly perturbed when the histidines were substituted by leucines (Figure 4.4). This is consistent with our hypothesis that at pH 6.8, the histidine residues are in their deprotonated forms, and thus demonstrating that leucine can act as a faithful representation of the constitutively deprotonated state.

When His_{72} and/or His_{75} were replaced with arginine, an amino acid that mimics a constitutively protonated histidine, the secondary structure of the variant $PRO^{PC1/3}$ decreases significantly, as seen from the changes in ellipticity at 222 and 208 nm (Figure 4.4). It is important to note two key points; first, the effects of the His_{72}Arg substitution alone are greater than the His_{75}Arg substitution alone, and second, the effects of the substitutions are cumulative, wherein there is a greater loss of secondary structure in the $His_{72/75}$Arg double mutant than in either of the single mutants. For comparison, the spectrum of WT-$PRO^{PC1/3}$ at pH 4.0 is shown (Figure 4.4, black dashed line). We see that the behavior of $H_{72/75}$R-$PRO^{PC1/3}$ at pH 6.8 recapitulates the behavior of the WT at pH 4.0, where we would expect both histidines to be protonated, and thus the propeptide to be maximally unfolded. With respect to the shifts in the troughs of the spectra around 208 nm, an indicator of alpha helicity, we see that the spectra of His_{75}Arg is slightly left-shifted towards 194 nm with respect to the wild type spectra, indicating a higher percentage of random coil. In contrast, the trough in His_{72}Arg is right-shifted, but significantly decreased in intensity, suggesting that the while the propeptide has more alpha helical content, it is losing overall structure; this perhaps reflects a transition from a more structured to a less

structured form, similar to what has been observed in subtilisin [66]. The cumulative effect of the arginine substitutions is seen in the spectra of the double mutant, where the spectral trough is roughly equivalent to that of WT-PRO$^{PC1/3}$ at pH 4.0.

The loss in structure is consistent with our earlier observations of the *in vitro* catalytic activities. Notably, since the double mutant ($His_{72/75}$Arg) is the only variant that exhibits constitutive activity at pH 7.4 (Figure 4.3A and B) these results strongly suggest that protonation of either His_{72} or His_{75} causes destabilization of the isolated PRO$^{PC1/3}$, however there is likely a requisite role for both residues in the pH-dependent activation of PC1/3.

A limitation of the above biophysical descriptions is the fact that they are done at static pH, whereas the PCs experience a broad range of pH as they transit the secretory pathway. Therefore, in order to further investigate the effect of these substitutions on the structural changes in the propeptide in a physiologically relevant way, we monitored unfolding of the isolated PRO$^{PC1/3}$ via changes in secondary structure between pH 8.0 and pH 4.0 (Figure 4.5). The WT-PRO$^{PC1/3}$ undergoes a cooperative sigmoidal transition from a well-structured state at pH 8.0 to a less structured state at pH 4.0 (Figure 4.5A and B, Black points). Data were fit to various models, including two-state unfolding (black line) and the Henderson-Hasselbalch equation accounting for either a single protonation event (green line), or two protonation events (red line). In each case, the midpoint of the transition between the well-structured state and the less

structured state occurred at ~pH 5.3, coincident with the pH of the later, more acidic compartment in which MAT$^{PC1/3}$ is activated in the secretory pathway.

Identical treatment of PRO$^{PC1/3}$ variants yield some interesting insights; at pHs above ~6.5, the leucine-variants of PRO$^{PC1/3}$ appear similarly well folded as the WT PRO$^{PC1/3}$, however as pH drops, the behavior of the mutants diverges. The H_{72}L-PRO$^{PC1/3}$ variant is largely stable, retaining its secondary structure at pH 5.0, and only undergoing insignificant unfolding (~20%, when compared with WT-PRO$^{PC1/3}$) as the pH approaches 4.0, potentially due to influence of secondary residues. Under identical conditions, the His_{75}Leu variant also appears unaffected by pH until pH 5.0, after which it loses about 60% of its secondary structure when compared with WT PRO$^{PC1/3}$ at pH 4.0. Thus, our results indicate that the overall pH-dependent stability of the His_{72}Leu variant is greater than that of His_{75}Leu as well as WT PRO$^{PC1/3}$, and suggest that the His_{72} residue contributes more towards the pH sensitivity. When the two mutations are combined, we observe that the variant PRO$^{PC1/3}$ is largely unaffected by the changing proton concentrations, reflecting the replacement of both histidines with nonprotonatable mimics. This suggests that while the effects of protonation and/or salt bridge formation on the structure of PRO$^{PC1/3}$ cannot be ruled out, the histidine protonation status is the primary driver of these structural changes.

In contrast, we note a striking effect of making the alternate substitutions, replacing histidines with positively charged arginine. Both the single His_{72}Arg and double $His_{72/75}$Arg exhibit significantly less secondary structure than WT-PRO$^{PC1/3}$ even at pH 8.0, and while the His_{72}Arg variant does undergo some

pH-dependent unfolding at pH ~6.0, again suggesting a minor role for protonation of secondary residues in the lower pH range, these changes are small. The His_{75}Arg exhibits the most striking of pH-dependent changes; at pH 8.0, the His_{75}Arg mutant has significant secondary structure, and appears to undergo a sigmoidal transition to a less folded state, with a midpoint of transition at pH ~7.0, again making this variant the most furin-like of those examined. Taken as a whole, these data again suggest a role for both His_{72} and His_{75}; in replacing both histidines with leucine, the pH-dependent loss of secondary structure can be abrogated, while the single substitutions merely attenuate this loss. Similarly, while replacing both residues with arginine renders $PRO^{PC1/3}$ likewise impervious to changes in pH, it exists in a baseline less-well folded state. The single arginine variants retain a degree of structural responsiveness to pH, undergoing structural transitions at distinct pHs, suggesting underlying differences in their pKas.

Effect of His substitution on the interactions of $PRO^{PC1/3}$ with $MAT^{PC1/3}$: While studies of the isolated $PRO^{PC1/3}$ are useful to understand the role of changing proton concentration on the conformational dynamics of this critical domain, ultimately $PRO^{PC1/3}$ exists as a complex with its cognate protease domain, and exerts its regulatory effects via inhibition. Therefore, we next determined the IC_{50} values of $PRO^{PC1/3}$ and its histidine variants for $MAT^{PC1/3}$ (Figure 6). The WT $PRO^{PC1/3}$ inhibits $MAT^{PC1/3}$ with an IC_{50} of 2.3 nM, consistent with previously published values (Figure 4.6A and B). Substituting either histidine singly or in combination with leucine did not significantly perturb the IC_{50} values of $PRO^{PC1/3}$ (Figure 4.6B) consistent with our observation that there are not

significant changes in secondary structure of these at pH 5.5 (Figure 3A). Likewise, consistent with our earlier observations, both the single $His_{75}Arg$ and the double $His_{72/75}Arg$ variants are poor inhibitors, with IC_{50} values approximately 10 to 15-fold higher than that of WT $PRO^{PC1/3}$. However, inhibition studies with $H_{72}R$-$PRO^{PC1/3}$ variant yielded some surprising results; unlike the other arginine variants, the IC_{50} of $H_{72}R$-$PRO^{PC1/3}$ was similar to that of the WT $PRO^{PC1/3}$ (Figure 4.6B), consistent with our prior observation that *in vitro* activity of this variant is only ~50% of the wild type. As we are yet unable to directly examine how pH affects the $PRO^{PC1/3}$:$MAT^{PC1/3}$ inhibition complex directly, due to the high concentrations of mature PC1/3 required to create stoichiometric complexes and difficulty in obtaining a truly catalytically inactive enzyme for obtaining stable complexes, we are left to speculate that perhaps there is interplay of additional residues at the interface of the protease and $PRO^{PC1/3}$ that allow the propeptide to remain tightly associated, despite having less structure and lower stability at acidic pH.

Discussion

Organellar pH is a critical regulator of a wide range of intracellular events, including zymogen activation, signaling cascades, ion channel activity, and receptor-ligand interactions. Each of these events must be tightly regulated both temporally and spatially in order to maintain organismal homeostasis, and many eukaryotic proteins have evolved to exploit environmental pH as a cue to regulate their activation via alterations in the protonation status of titratable amino

acid side chains. Not surprisingly then, protonation with increasing organellar acidification is the major biochemical cue that regulates the final activation step of PCs. We are only just beginning to understand how the differing pH sensitivity of individual PCs is encoded, and how the addition of a single proton can drive the biochemical and biophysical changes required to initiate protease activation. We have earlier demonstrated that protonation of His_{69} can disrupt a hydrophobic pocket within PRO^{FUR} to drive local unfolding of the secondary cleavage site loop to facilitate PRO^{FUR} proteolysis that results in furin activation [103, 104]. While His_{69} is conserved within all PCs, individual PCs nonetheless undergo activation at different pHs, with actual proton concentrations that vary over ten-fold, suggesting that additional complexities remain to be determined.

Our results show that while histidine residues are mostly conserved within orthologues of PC1/3 and furin individually, they show significant differences between the two families, with only the primary pH sensor residue (His_{69} in PRO^{FUR}, His_{72} in $PRO^{PC1/3}$) absolutely conserved throughout the alignment (Figure 4.2). The remaining histidines, His_{67}, His_{75} and His_{85} in $PRO^{PC1/3}$ are located within proximity of the cleavage site loop. His_{67} (which corresponds to His_{66} in furin) is situated at the interface with the protease domain and has been demonstrated to have no influence on pH sensing in furin [103]. Furthermore, preliminary studies using leucine and arginine substitutions at positions 67 and 85 in PC1/3 suggest that these variants do not affect the secondary structure of the isolated propeptide domain, or their IC_{50} values (data not shown). It is noteworthy that (i) the counterpart to His_{69}, which is solvent accessible in furin,

appears more buried within PC1/3; (ii) the imidazole side chain of His_{75} is solvent exposed and does not have an analogous counterpart in furin; and (iii) His_{75} precedes a proline residue in PC1/3. Experimental studies have established that when a histidine side-chain precedes a proline residue, the proline isomerization rates increase up to 10-fold when the imidazole side-chain is protonated relative to the deprotonated state [212, 213]. In this manuscript, we therefore analyzed the role of the conserved histidine in PC1/3 (His_{72}) as well as that of a second histidine (His_{75}), which we propose acts as a 'gatekeeper' that regulates solvent accessibility, and thus protonation, of the primary pH sensing histidine.

To decipher the role of the two histidine residues in regulating the pH-dependent activation of PC1/3, we used constructs with a C-terminal KDEL sequence that retains the noncovalently bound inhibition complex within the ER; in this case, proPC1/3 has undergone the pH-independent primary processing subsequent to folding, yet the propeptide remains associated, acting as an inhibitor of protease activity. Our results indicated that mutation of either His_{72} or His_{75} to a nonprotonatable mimic was sufficient to block the second processing event, and thus activation of PC1/3, at all pH's assayed, indicating that the protonation of both histidine residues are necessary and neither alone are sufficient to drive pH dependent activation of PC1/3. It is noteworthy that when replaced with an arginine, a surrogate that mimics a constitutively protonated state, the H_{72}R-PRO$^{PC1/3}$ does not promote activation while the H_{75}R-PRO$^{PC1/3}$ undergoes marginal activation at pH 7.4. Although activation at pH 6.4 and 5.4 increases for both H_{72}R-PRO$^{PC1/3}$ and H_{75}R-PRO$^{PC1/3}$, only the double variant

($H_{72/75}$R-$PRO^{PC1/3}$) mediates robust activation at all pHs. On this basis, we propose that both His_{72} and His_{75} appear to work cooperatively to mediate the pH-dependent activation of PC1/3 (Figure 4.3).

The biophysical analysis of the isolated propeptides further supported the idea that both histidine residues were important in the mechanism of activation for this protease. Here, the pH-dependent structural transitions seen in the various mutant propeptides are particularly insightful. The H_{72}L-$PRO^{PC1/3}$, H_{75}L-$PRO^{PC1/3}$ and the $H_{72/75}$L-$PRO^{PC1/3}$ variants behave similar to WT $PRO^{PC1/3}$ between the pH of 6.0 to 8.0. However, at pH less than 6.0, the WT $PRO^{PC1/3}$ begins to unfold, with a calculated pKa of ~5.2 when the data is fit to a two-proton model (Figure 4.5A). Under these conditions, all histidine to leucine $PRO^{PC1/3}$ variants are significantly more stable than the WT $PRO^{PC1/3}$ and although the H_{75}L-$PRO^{PC1/3}$ loses about 50% of its ellipticity at 222nm when compared to WT $PRO^{PC1/3}$, H_{72}L-$PRO^{PC1/3}$ and $H_{72/75}$L-$PRO^{PC1/3}$ variants undergo only marginal unfolding (~20%) under identical conditions (Figure 5B). While these results indicate that a single substitution of histidine with nonprotonatable leucine residues at position 72 or 75 blunts the ability of $PRO^{PC1/3}$ variant to respond to more acidic pH, the subtle differences in the behavior of the individual variants allows one to begin teasing apart the differing roles of these residues.

First, given that the single H_{72}L variant is sufficient to attenuate the pH-dependent structural changes within the propeptide (Figure 4.5B) and block pH dependent activation (Figure 4.3A) we propose that the protonation status of His_{72} is a major driver of the requisite unfolding of the cleavage loop of $PRO^{PC1/3}$.

Second, the observation that the H_{75}L undergoes partial unfolding, albeit to a lesser degree than the WT $PRO^{PC1/3}$, indicates that while protonation of His_{72} is necessary, it is not sufficient, and thus the combinatorial effect of protonation of both residues appears to be necessary for activation of PC1/3. As a final note, the small change in secondary structural content in the double $H_{72/75}$R-$PRO^{PC1/3}$ variant is likely due to influence of other titratable groups, which at low pHs may enhance the local changes required for propeptide processing and release, and thus further studies are needed to more fully investigate this possibility.

By introducing an arginine in place of the histidines of interest, we were able to assess the effect of a positive charge at these positions within the propeptides. Again, the differences between the variants sheds light on the differing role that each of these histidines are playing, and lends further support to the notion that there is a synergistic effect between His_{72} and His_{75} that allows for the precise spatiotemporal regulation of the loosening of secondary structure required for $PRO^{PC1/3}$ processing. Even at pH 8, H_{72}R-$PRO^{PC1/3}$ displays markedly less secondary structure than the WT-$PRO^{PC1/3}$, reflecting its role as a major driver of destabilization. There is an additional loss of alpha helicity as pH drops below ~6, either due to the aggregate effect of protonation of His_{75} or the influence of other side chains. The behavior of H_{75}R-$PRO^{PC1/3}$ is again of particular interest; while somewhat destabilized relative to WT-$PRO^{PC1/3}$ it is nonetheless well structured at pH 8 when compared with H_{72}R-$PRO^{PC1/3}$ and $H_{72/75}$R-$PRO^{PC1/3}$. As pH is lowered, H_{75}R-$PRO^{PC1/3}$ immediately begins to lose secondary structure, but plateaus at pH 6. This again highlights our assertion that

the positive charge at either of these positions singly is not sufficient to account for the activation behavior of PC1/3, but rather that cooperative action between early protonation of His_{75} followed by later protonation of His_{72} is required.

A final piece of insight is offered by the ability of the propeptides to inhibit the protease domain (Figure 4.6). As would be expected, IC_{50} values (at pH 6.8) for the leucine variant propeptides roughly approximate those of the wild-type, while those for the His_{75}Arg and $His_{72/75}$Arg variant are ~15-fold higher, indicating that they are notably weaker inhibitors. At first it may seem surprising that the His_{72}Arg variant has an IC_{50} that is comparable to that of the wild-type, however we propose this reflects a critical role of the histidine at position 75; in addition to modulating the solvent accessibility of the hydrophobic pocket it overlays, protonation of His_{75} may be involved in a destabilization of the propeptide: protease complex. As we are yet unable to produce the mature protease in sufficient quantities, we are unable to more directly determine k_{on}/k_{off} and more thoroughly characterize the interaction between the protease and its cognate propeptide.

Taking the above as a whole, we can then propose a mechanism by which PC1/3 is activated (Figure 4.7). We believe that the two histidine residues within the cleavage loop are in proximity of one another, and require sequential protonation to allow for activation. His_{75}, which sits at the top of the hydrophobic pocket, is likely protonated first, and begins the local unfolding of the cleavage loop. It is tempting to speculate that protonation of His_{75} facilitates local unfolding by enhancing the kinetics of cis-trans isomerization of Pro_{76} as demonstrated by

systematic studies in a pentapeptide series Ac-Ala-Xaa-Pro-Ala-Lys-NH2 [212]. These studies, which examined the influence of each of the 20 amino acids at position Xaa on the energetics of proline isomerization using NMR spectroscopy, demonstrated that the rates of proline isomerization are enhanced several fold only when the side-chains of tyrosine and histidine residues are protonated. The occurrence of histidine residues preceding a proline is higher than the overall frequency of the individual amino acids would predict [214]. Thus we propose that the protonation of His_{75} enhances local unfolding of the loop, which allows His_{72} to become more solvent accessible, and simultaneously disrupt the interface between the propeptide and protease. Once solvent accessible, His_{72} can then be protonated, thus delivering the second and final blow that drives the processing and dissociation of the propeptide, thus releasing inhibition from the protease. In this model, His_{75} acts as a gatekeeper residue, blocking access to the hydrophobic pocket in its deprotonated state, a restriction that can be removed via protonation. The model for two protonation events is also supported by CD data on the WT-$PRO^{PC1/3}$. Notably, when we fit the CD data from pH-dependent unfolding experiments to a modified form of the Henderson-Hasselbalch equation [211], we find that a model of two titratable groups better fit the conformational changes in $PRO^{PC1/3}$ in response to changing pH than one that only allowed for a single protonation (Figure 4.5A). We also speculate the close proximity of His_{72} and His_{75} may also be critical to understanding the mechanism of activation, and reason that PC1/3 is only activated in acidic environs. When both residues are unprotonated, hydrophobic packing of the core

of the propeptide is the major stabilizing force, similar to the furin propeptide; one must ask then, is it simply a greater degree of stability in the structure of $PRO^{PC1/3}$ that requires the stronger disrupting force of two positive charges to drive its unfolding, or does the protonation of one residue influence the protonation of the second? Concomitant work in our laboratory has been focused on defining pKa values of the histidines within PRO^{FUR} and $PRO^{PC1/3}$. While the pKa of His_{67}, His_{75} and His_{85} are all ~6.0, the pKa of His_{72} is acid-shifted, with a pKa of ~5.5 [106], a value that similar with the pH of activation of PC1/3. Therefore, we believe that the spatial arrangement and interaction of His_{72} and His_{75} can explain why PC1/3 is activated by a ten-fold higher proton concentration than its paralogues furin.

This work offers significant insight into the question of how each member of the PC family encodes unique organelle-specific information about activation and processing. We believe that while the primary pH sensor (His_{69} in furin, His_{72} in PC1/3) is evolutionarily conserved throughout PCs, concomitant with histidine enrichment has been a tuning of the pH sensitivity of these proteases. As seen in the multiple sequence alignment, only the position of the primary pH sensor has been absolutely conserved, thus domain-specific conservation, as opposed to position-specific conservation, seems to be crucial to encoding pH sensitivity in the propeptides of proteases in this family. Looking more broadly at secreted proteases across eukaryotes and prokaryotes, we had previously reported that there is an enrichment of histidine residues in the propeptides of secreted eukaryotic proteases that is not seen in their prokaryotic paralogues. Again, we

see that while there is a bias for more histidines within this particular domain, there is no appreciable bias for specific positions [205]. These observations highlight the importance of spatial juxtapositioning of titratable groups in regions of proteins in order to tune their pH sensitivity, an evolutionary adaptation that was likely concomitant with the emergence of compartmentalization and specialization allowed by the secretory pathway within eukaryotic cells.

The issue of “mistuning” via mutation is also interesting to consider. The most obvious case to consider would be the loss of the conserved pH sensor, yielding either an inactivatable or constitutively active protease, depending on the type of mutation. More interesting however, would be the mutation of one of the surrounding residues, and its effect on the local environs and protonatability of not only the primary pH sensor, but also the gatekeeper or tuning residues. Evidence for the plausibility of this notion is lent by several single nucleotide polymorphisms in PRO^{FUR} and $PRO^{PC1/3}$ that have been associated with various cancers [including head and neck carcinoma, lung adenocarcinoma [215]] or obesity [PC1/3 ΔR80Q [113]]. Further characterization of these polymorphisms would be insightful not only in understanding the basis on which they are able to cause disease, but also in the broader sense of understanding what role propeptides have in dictating proteolytic activity and processing. Similarly, environmental homeostasis is required for optimal proteolytic function; perturbation of cellular pH through disease [75] may cause activation of a PC prematurely, or prevent activation all together. The role of the pH gradient in the secretory pathway in orchestrating proteolytic processing of substrates cannot be

understated, thus as we continue to understand the interplay between the proprotein convertases, proton concentration, and pH sensors, we hope to gain better insight into how things go wrong in the disease-state, and how we can better approach solutions to restore homeostasis.

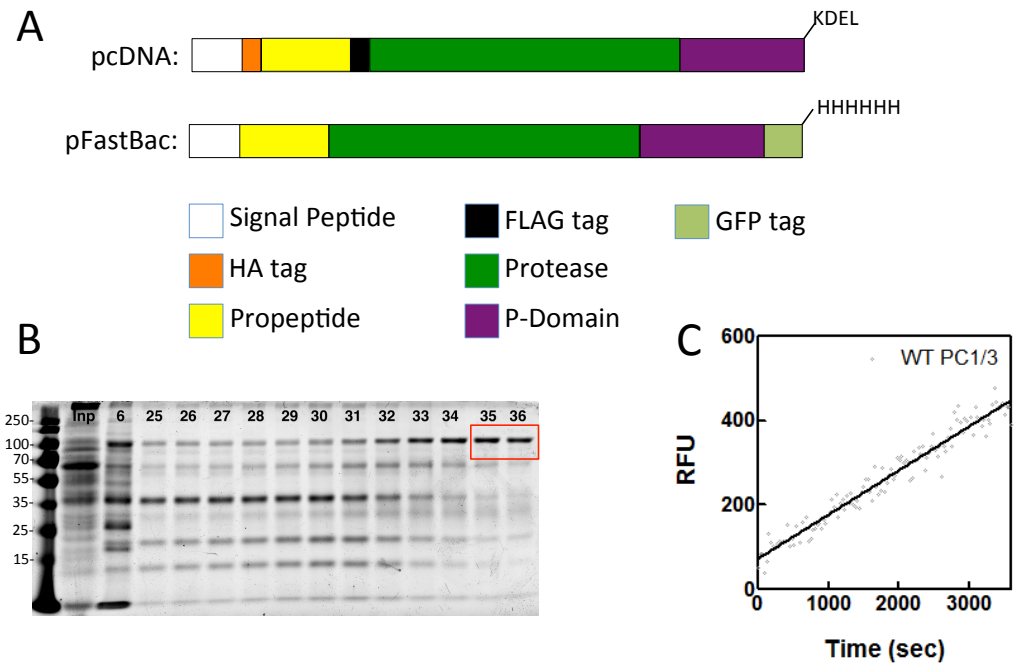


Figure 4.1: Constructs and Characterization of Baculovirus-Produced PC1/3

A) Schematic representation of the constructs used in this manuscript. Top; insert cloned into pcDNA3.1, where the coding sequence for WT mouse PC1/3 has been truncated at Arg_{618} with the addition of a –KDEL sequence, an HA tag inserted at the N-terminus of the propeptide, and a FLAG-tag inserted at the N-terminus of the protease domain. Bottom; the coding sequence for WT mouse PC1/3 was truncated at Arg_{618} and a C-terminal GFP-tag and 6x-His tag added for insertion into the pFastBac vector for baculovirus production. **B**) Representative SDS-PAGE gel showing input loaded (Inp) and fractions collected following elution from Ni-NTA column (6, 25-36). Fractions indicated in the red box were pooled, concentrated, and used for assays. **C**) Representative enzyme progress curve showing baculovirus produced WT PC1/3 processing of the fluorogenic substrate. The curve suggests that PC1/3 has fully matured and displays normal enzyme characteristics.

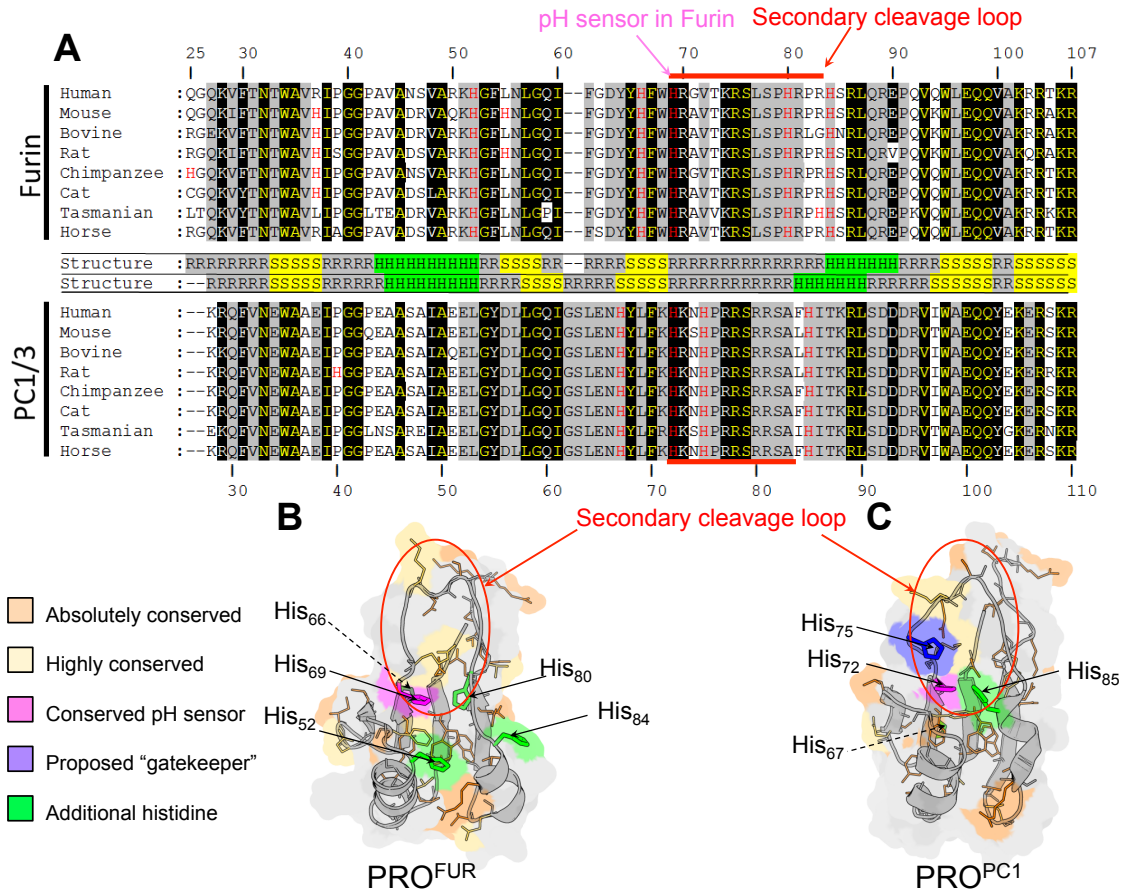


Figure 4.2: Comparative analysis of PRO^{FUR} and PRO$^{PC1/3}$
A) Multiple sequence alignment of eight orthologues of PRO^{FUR} (top) and PRO$^{PC1/3}$. Names of the species are denoted on the figure and their corresponding accession numbers for furin are; Human: P09958, Mouse:P23188, Bovine: Q28193, Rat: P23377, Chimpanzee: H2QA34, Cat: M3W594, Tasmanian Devil: G3W614; and Horse: F7DHR9. Similarly, the species and their corresponding accession numbers for PC1/3 are as follows; Human: P29120; Mouse: P63239; Bovine: Q9GLR1; Rat: P28840; Chimpanzee: H2QR92; Cat: M3W5Q0; Tasmanian devil: G3VJH4; and Horse: F6TCF0**.** Residues are numbered with respect to the human homologue of each peptide. Residues that are absolutely conserved in orthologues of either PRO^{FUR} or PRO$^{PC1/3}$ are shaded in black, while those residues conserved in orthologues of both PRO^{FUR} and PRO$^{PC1/3}$ are in yellow text with black shading. Highly conserved residues are shaded in grey. Histidine residues are highlighted in red text, with the conserved pH sensor indicated by the magenta arrow. Residues comprising the secondary cleavage loop are indicated by the red bar. In between sequence alignments, secondary structures are indicated, where R indicates random coil, S indicates beta sheet, and H indicates alpha helix. **B**) Homology model for PRO^{FUR} [92] and **C**) Solution structure (1KN6) of PRO$^{PC1/3}$. The absolutely conserved residues are in B and C are colored orange, while highly conserved residues are colored yellow. The conserved pH sensing histidines in PRO^{FUR} and PRO$^{PC1/3}$ are colored magenta, while additional histidines are colored green, with the proposed "gatekeeper" histidine residue in PRO$^{PC1/3}$ is colored purple. The secondary cleavage loop is circled in red.

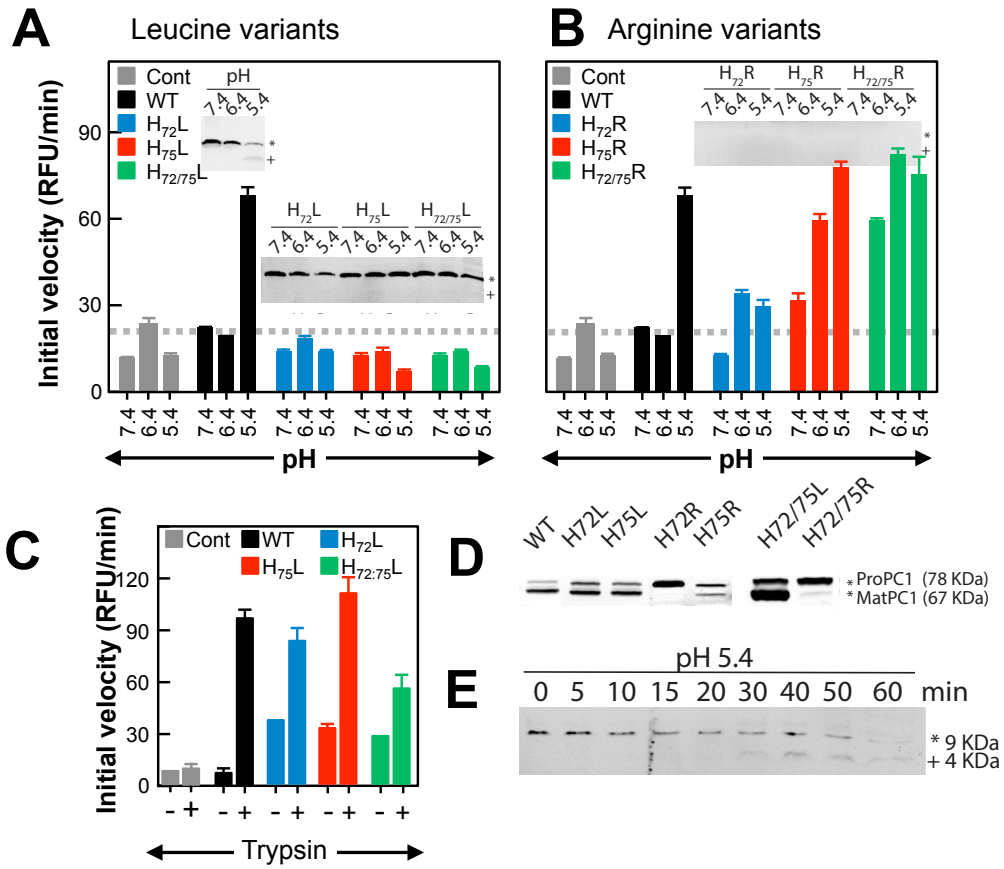


Figure 4.3: pH-dependent activation of the ER retained precursor variants of PC1/3.

Activity from ER extracts of cells transfected with KDEL-tagged proPC1/3 after preincubation at indicated pH. The activity profiles for WT precursor activated using various pH are shown in black bars, variants at position 72 in blue, position 75 in red, and at both 72 and 75 in green. Histidines at these positions were mutated to leucine (**A**) or arginine (**B**). Cells transfected with empty vector (Cont) were treated identically, with the results shown by grey bars. Grey dashed lines indicate the threshold for baseline activity, below which an enzyme is considered inactive. Results are the mean ± SD of three independent experiments performed in triplicate. Insets show western blot detecting the product of the primary cleavage, which results in a 9kDa fragment of $PRO^{PC1/3}$; Asterisk and plus indicate the uncleaved (~9kDa) and cleaved (~4kDa) of WT $PRO^{PC1/3}$ along with the leucine and arginine variants of His_{72} and His_{75}. **C**) Enzymatic activity using ER extracts from cells transfected with WT $PRO^{PC1/3}$, along with $H_{72}L$-$PRO^{PC1/3}$, $H_{75}L$-$PRO^{PC1/3}$ and $H_{72/75}L$-$PRO^{PC1/3}$ after preincubation at pH 7.4, followed by incubation in presence or absence of trypsin. **D**) Western blot showing the uncleaved precursor (proPC1/3; 78 kDa) and the mature protease domain ($MAT^{PC1/3}$; 67 kDa) for the WT and histidine variants. **E**). Time course of processing of the WT $PRO^{PC1/3}$ after incubation at pH 5.4 for the indicated number of minutes. Asterisk and plus indicate the uncleaved (~9kDa) and cleaved (~4kDa) forms of $PRO^{PC1/3}$.

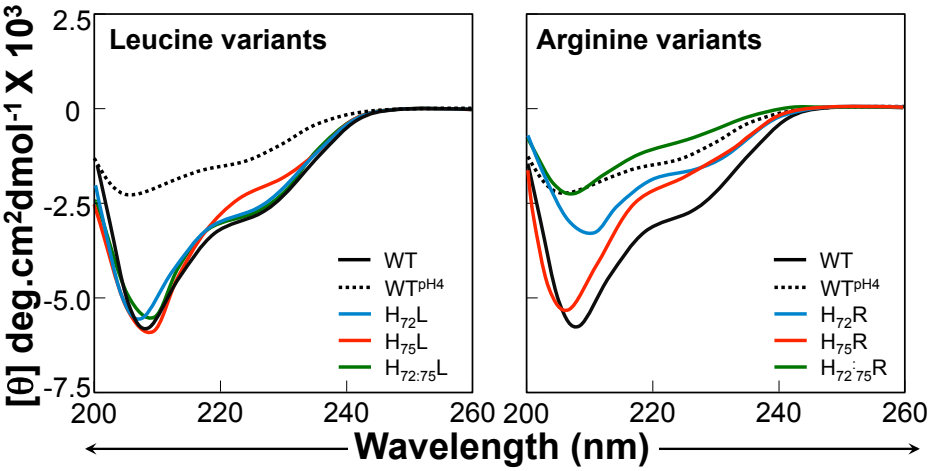


Figure 4.4: Effect of His substitution on the secondary structure of isolated PRO$^{PC1/3}$

Circular dichroism spectra of histidine to leucine **(left)** and histidine to arginine **(right)** of isolated PRO$^{PC1/3}$ at pH 6.8 and plotted as molar ellipticity (θ) deg.cm^{2}dmol^{-1}. Spectra for the wild type are shown in black line (pH 6.8) or dashes (pH 4.0), mutations at position 72 in blue, mutations at position 75 in red, and double mutants in green.

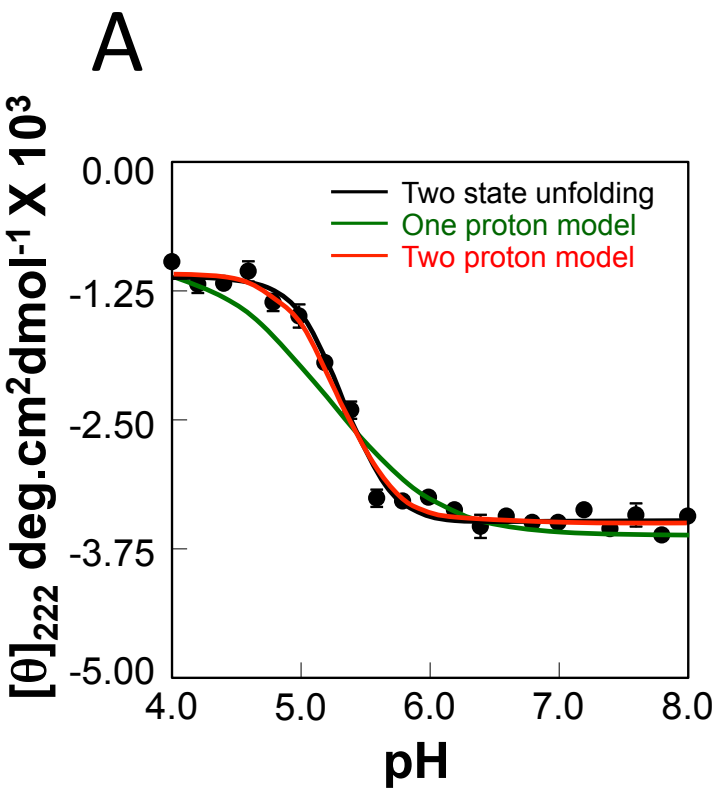
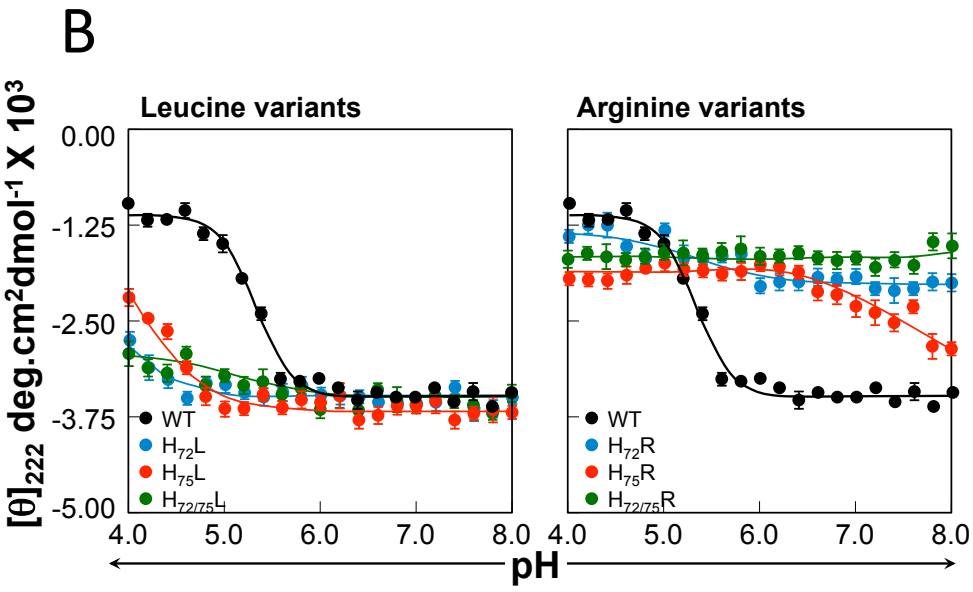


Figure 4.5: Effect of pH on secondary structure of isolated IMC$^{PC1/3}$
A) pH induced unfolding of the isolated WT-PRO$^{PC1/3}$ (black dots). Unfolding was monitored by changes in ellipticity at 222nm as a function of pH, and plotted as molar ellipticity (θ) deg.cm^{2}dmol^{-1}. Data was fit to equations describing two-state unfolding (black line), or the Henderson-Hasselbalch equation for one proton (green line) or two protons (red line).



B) pH induced unfolding of histidine to leucine (left) and histidine to arginine (right) variants compared with the WT $PRO^{PC1/3}$ (black), with mutations at position 72 shown in blue, mutations at position 75 in red and double mutants in green.

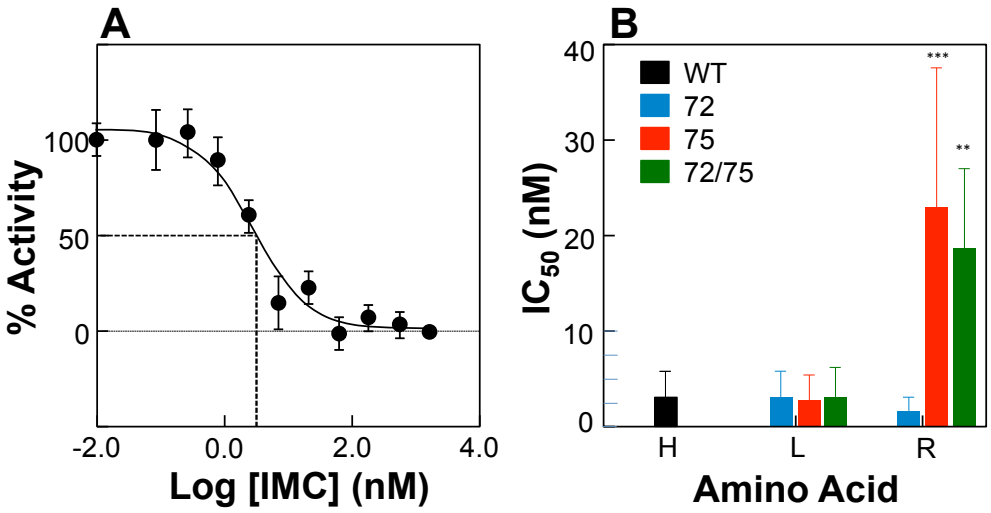


Figure 4.6: IC_{50} values of WT $PRO^{PC1/3}$ and its variants for interactions with $MAT^{PC1/3}$.
A) Representative plot of activity of WT $MAT^{PC1/3}$ as a function of concentration of WT-$PRO^{PC1/3}$. Values are mean ± SD of one experiment performed in triplicate. **B**) IC_{50} values determined for WT $PRO^{PC1/3}$ and the histidine variants at pH 6.8, plotted as the mean ± SD of three independent experiments performed in triplicate, and those values that are statistically different than the WT are indicated by asterisk (*** ~ $p<0.001$; ** ~ $p<0.01$).

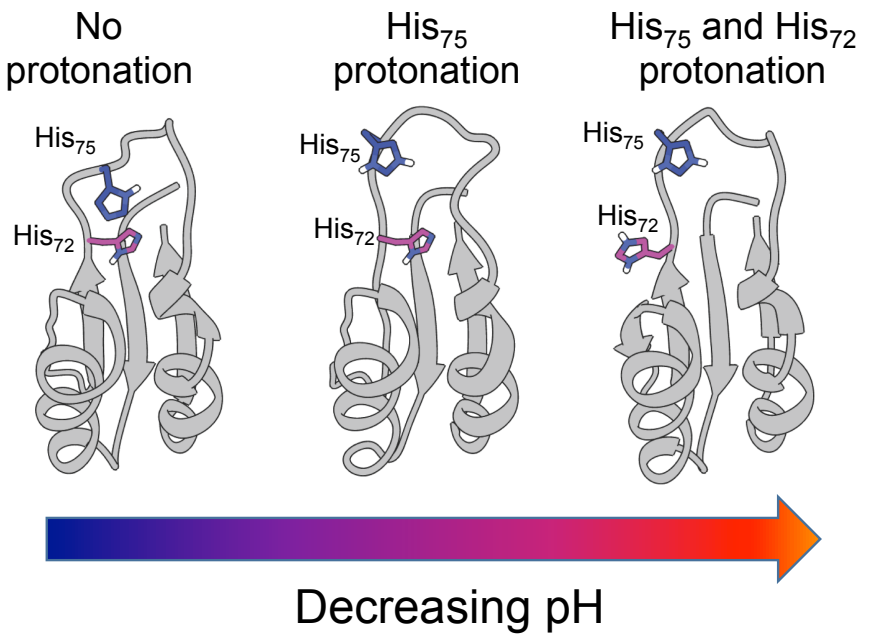


Figure 4.7: Model of histidine protonation in PRO$^{PC1/3.}$
Two-dimensional cartoon representations of the NMR solution structures of the PC1/3 propeptide (1KN6) displaying histidines side chains of interest. Left: Both histidines are deprotonated, maintaining packing of core and protecting cleavage site. Center: His_{75} protonated, causing partial unfolding of the cleavage loop, and exposing $His_{72.}$ Right: Both histidines protonated, causing cleavage loop to unfold completely, and exposing cleavage site to allow processing.

Chapter V: Protein memory: Single nucleotide polymorphisms in the propeptide of PC1/3 can alter protease function

Danielle M. Williamson and Ujwal Shinde

This chapter is currently in preparation for publication. Experiments and manuscript preparation have been carried out by DMW

Introduction

The conversion of precursor proproteins to their active forms via endoproteolytic cleavage is a conserved theme in the biosynthesis of biologically active proteins [13]. The proprotein convertases (PCs) represent a conserved family of eukaryotic proteases that are responsible for the processing of a diverse pool of secreted proproteins to their active forms [115]. Comprised of nine serine endoproteases (PC1/3, PC2, furin, PC4, PACE4, PC5/6, PC7/LCP/PC8, SKI/S1P, and NARC1/PCSK9), the PCs play a central role in the maintenance of physiologic homeostasis within cells and tissues [122, 161, 183]. In the nearly thirty years since their initial description, myriad studies have elucidated the multifaceted role of these proteases, and highlighted their importance not only in prohormone processing, but also embryogenesis, tissue remodeling, cholesterol homeostasis, as well as activation of bacterial toxins and processing of viral envelope glycoproteins [122, 183]. Consistent with this, mutations in PCs have been associated with pathologies such as obesity, hypercholesterolemia and carcinoma [183, 216, 217].

Structurally, the PCs are related to the bacterial subtilisins and yeast kexin, and as such, share a set conserved domains that guide their function. The core domains are i) a signal peptide that targets the nascent protein to the endoplasmic reticulum (ER) and thus the secretory pathway, ii) a propeptide critical for protein folding and regulation of activation, and iii) a catalytic domain possessing the catalytic triad of an aspartate, histidine, and serine residue;

additionally, the PCs possess two additional domains, a P-domain and a C-terminal tail, which have been implicated in regulation of calcium dependence, localization, and catalytic activity [122].

While not a part of the mature, functional protease, the propeptide nonetheless plays a critical role in the functioning of the proprotein convertases. Not only does it regulate spatiotemporal activation [92, 103, 105], it acts as a single-turnover catalyst of folding, and thus can be considered an intramolecular chaperone (IMC) [88, 89]. Based on extensive characterization of the subtilisins, the propeptide serves as a structural scaffold that facilitates folding of the catalytic domain. Upon completion of folding of the protease domain, the peptide bond joining the propeptide to the protease is cleaved in an autocatalytic process. However, the propeptide remains associated with the mature protease in an inhibition complex and it is not until the complex is properly trafficked to the appropriate organelle that a second autocatalytic event at an internal site within the propeptide triggers the release and subsequent activation of the protease. We have previously demonstrated that this second autoprocessing event is pH-dependent, and is sensed via protonation of a conserved histidine residue within the propeptide [103] [105] that drives local structural changes that expose the cleavage site [104].

As inappropriate regulation of proteolytic activity can have devastating consequences on organismal homeostasis [112, 204], it is tempting to speculate that mutations in the propeptide domain of the PCs could alter folding and/or activation of the protease domain, thus causing downstream effects in substrate

processing. Specifically, mutations in the propeptide can be imprinted onto the fold of the protease domain, even when then the primary sequences of the functional domains are identical. This phenomenon was originally described in subtilisin, where point mutations in the propeptide resulted in folded proteins differing in activity, substrate specificity, and selectivity of inhibition [48, 112]. Additionally, mutations in residues surrounding the cleavage site or pH-sensor within the propeptide could alter autoprocessing and activation, such as those noted in the propeptide of PC1/3 [113, 218, 219].

PC1/3, a neuroendocrine member of the PC family, is responsible for the processing of a number of important proproteins to growth and energy balance, including proinsulin, proglucagon, proopiomelanocortin, and proghrelin, among others [155, 220-223]. Given the range and critical role of these hormones, it is not surprising that deficiencies in PC1/3 frequently lead to complex metabolic phenotypes, which have been reported both in mouse models and in humans. To date, four patients have been identified with mutations in PCSK1 gene that result in PC1/3 deficiency, and cause early-onset obesity, hyperphagia, hypoadrenalism, reactive hypoglycemia, malabsorbtive diarrhea, hypogonadism, and diabetes insipidus [199, 201, 202, 224]. A number of single nucleotide polymorphisms in the PCSK1 gene have likewise been identified and described; these polymorphisms affect catalytic activity and substrate processing, and have been associated with increased risk of diabetes and obesity (19-22), as well as premature ovarian failure [225] and coronary artery disease [226]. In many cases, even minor alterations in PC1/3 activity can shift the homeostatic balance

of an organism, predisposing them to metabolic derangements [126, 224]. Not surprisingly, PCSK1 has been called the third most important gene contributing to extreme obesity [197]. While the PC1/3-null mouse does not phenocopy the human [194, 227], a missense mutation in PCSK1 ($N_{222}D$) results in a hypomorph of PC1/3 that causes obesity and hyperphagia [228].

Of particular interest in our research is the role of the three single nucleotide polymorphisms within the propeptide region of the *PCSK1* gene that result in missense variants (Figure 5.1)[229]. While several SNPs in the region of *PCSK1* encoding the protease domain have been characterized, only one of SNPs within the propeptide has been evaluated [113]; furthermore, a biochemical understanding of how these polymorphisms affect the function of PC1/3 remains elusive. As noted above, as the propeptide acts as both a chaperone for folding as well as an inhibitor responsible for regulating activation, it is likely that the variant propeptides affect structural heterogeneity and/or altered pH sensing despite the sequence homogeneity of the protease. These changes would undoubtedly have significant effects on protease function and substrate specificity that may have marked consequences for cellular and organismal homeostasis. Therefore, we chose to undertake biophysical and biochemical studies to lend further insight into the impact of these SNPs on protease function. Our data demonstrates that the PC1/3 protease chaperoned by SNP-variant propeptides display biochemical differences from a seemingly identical protease chaperoned by the wild type propeptides. These differences in catalytic activity and interactions have the potential to cause imbalances in prohormone

processing that underlie the complex metabolic phenotypes of type 2 diabetes and obesity.

Results

As aforementioned, it is likely that mutations within the propeptide of PC1/3 could alter the structure of the propeptide or its interaction with its cognate protease, thus impacting it's dual roles as a chaperone and inhibitor. Support for this hypothesis comes from work by Rabah et al. [218], who demonstrated that mutation of a single amino acid could alter the inhibitory behavior of propeptides toward their cognate protease. Additionally, Pickett, et al. [113], showed that when Arg_{80} was replaced by Gln, the most commonly occurring polymorphism within the PC1/3 propeptide, the maturation and *in vitro* catalytic activity of PC1/3 was affected. Given this insightful data, we first asked whether there were additional polymorphisms within *PCSK1* that encoded missense mutations in the propeptide of PC1/3. Data from the dbSNP [229] was compiled, which identified three single nucleotide polymorphisms, including the previously described $R_{80}Q$ (Figure 5.1). All three are in close proximity of the secondary cleavage site (R_{81}), and are predicted to be deleterious to protein structure and function by the PolyPhen and SIFT tools, which use available sequence, phylogenetic and structural information to predict the impact of amino acid substitutions [230, 231]. Therefore, we chose to analyze the two novel variants, as well as the aforementioned $R_{80}Q$, with a focus on the propeptide with the goal of providing insights into the critical role of this domain in convertase function from a

biochemical and biophysical perspective, and the implication of these variants on human health and disease.

Structure of isolated propeptides

As the structure of the propeptide is central to both its role as a chaperone as well as an inhibitor, we began by analyzing the secondary structure of the isolated propeptides (Figure 5.2). The spectrum of the wild type propeptide indicates a well-structured state. Relative to the wild type, we see that two of the SNP-variant propeptides ($P_{76}S$ and $R_{80}Q$) are less well structured, and one ($R_{78}S$) is more structured. It is notable that the spectra of the $R_{78}S$ and $R_{80}Q$ variants have a spectral minima at ~208 nm, like the wild type, indicating that they have similar but not identical secondary structures, representing homogenous but structurally distinct folded protein populations. While CD spectroscopy cannot provide definitive identification of where specific elements of secondary structure are, or how they are changing, the spectra do indicate that there is an overall loss of secondary structure as a result of the variants. In contrast, the minimum of the $P_{76}S$ variant is left-shifted toward 194 nm, indicating a greater proportion of random coil.

Inhibitory capabilities

It has been previously demonstrated that a furin propeptide with increased secondary structure within the isolated propeptide correlates with an increased affinity for its cognate catalytic domain [104]. Given the differences observed within the secondary structures of the SNP-variant propeptides, the inhibitory role

of the propeptide was therefore of particular interest. To test this, we first determined IC_{50} values of the SNP-variant propeptides for wild type PC1/3 (i.e.: protease produced with the wild type propeptide) (Figure 5.3). Concordant with our observation that the $P_{76}S$ and $R_{80}Q$ are less well structured than the wild type propeptides, we see that the IC_{50} values of these propeptides are significantly higher than that of the wild type. The IC_{50} of the $R_{78}S$ variant, which is more well structured than the wild type, is similar to that of the wild type, suggesting that there may be additional factors contributing to the inhibitory capability of the propeptide, including alterations in the efficiency of propeptide processing.

We were also curious if the variant propeptides demonstrate a higher affinity for their cognate protease domains, based on prior work in the bacterial subtilisins [112]. We discovered that the $P_{76}S$ and $R_{80}Q$ variants were ~4- and 10- fold better inhibitors of their cognate proteases, respectively, than the wild type protease. Interestingly, the $R_{78}S$ variant was a poorer inhibitor of its cognate protease, requiring ~14-fold higher concentration of propeptide to achieve a 50% reduction in activity. A final point worth noting is that in all cases, i) the protease domain does not have any mutation and (ii) the variant propeptides display lower affinity for their cognate protease than the wild type propeptide for the wild type PC1/3.

Processing of synthetic substrates

The fact that we see differences in the IC_{50} values for the variant propeptide towards the wild type-chaperoned protease compared to the variant-chaperoned protease suggested that there are differences in the proteases

themselves, even after the propeptide has been released and degraded. Therefore, we compared Km of the wild type PC1/3 for two synthetic substrates to those of the SNP-variant chaperoned PC1/3 (Figure 5.4). The first substrate contained the RTKR consensus furin cleavage sequence, commonly used in enzymatic assays of the proprotein convertases, while the second contained a variant RTAR motif. The Km of the alternate substrate for the wild type protease was ~10 higher than that of the canonical substrate, reflecting the loss of the dibasic consensus motif. Enzymatic activity of the variant PC1/3 proteases was altered against the RTKR as well as the RTAR substrate; Kms of the RTKR substrate were higher for all of the variant PC1/3 proteases than the wild type, suggesting catalytic differences, despite their primary sequence homology. Additionally, the differences in the Km values for the RTAR substrate revealed interesting potential implications for cleavage site preference in the variant proteases. The $R_{78}S$ was most similar to the wild type, with a ~10-fold higher specificity for RTKR over RTAR, and while RTKR is a better substrate for the $R_{80}Q$ variant, the difference in Kms is only ~3-fold, suggesting a loss of stringency. The $P_{76}S$ variant yielded the most interesting result, displaying a preference for the variant substrate lacking the dibasic motif over the canonical cleavage sequence.

Materials and methods

Baculovirus Production and Expression of Protein

Enzymatically active, mature PC1 was expressed via the baculovirus system, as previously reported [105]. Briefly, baculovirus expressing either the wild type PC1/3 or PC1/3 with the propeptide polymorphisms noted was used to infect Sf9 cells at a MOI of 2 in SF900-III media. After 24h, the media was clarified by two subsequent centrifugation steps at 500x and 10,000x g, respectively, for 20 minutes at 4C. Protein was precipitated using 20% PEG-8000 and resuspended in 100-fold smaller volume of 50 mM MOPS at pH 6.5,300 mM NaCl, 5mM CaCl2, and 0.01% Brij-35 (Buffer A) mixed with fresh cOmplete EDTA-free Protease Inhibitor cocktail (Roche Diagnostics GmbH). After clarification by centrifugation at 10,000g for 20 minutes, 10 mM imidazole was added and the buffer was agitated for 2h with 1 ml Ni-NTA resin (Thermo scientific). Afterwards the buffer was packed into a column and extensively washed (>20x column volumes) of Buffer A with 25 mM imidazole. PC1 was then eluted using a linear imidazole gradient to 500 mM and fractions containing GFP fluorescence and PC1 activity were pooled, flash-frozen, and stored at -80C.

Expression and Purification of Propeptide

Codon-optimized sequences encoding either the WT PC1/3 propeptide or where the single nucleotide substitutions had been made to encode the polymorphisms noted were cloned into the pET11b backbone and expressed in BL21(DE3) E. coli as previously described [104, 105]. Protein was purified from

the soluble fraction by ion exchange after cell lysis via French pressure cell, and dialyzed into 6M guanidinium HCl containing buffer [50 mM Tris-HCl, 6M guanidinium HCl, pH 6.5] for long-term storage. Before use, proteins were refolded by dialysis against refolding buffer [10mM Tris, 10mM Cacodylate, 10mM NaOAc, 150mM KCl, 5 mM $CaCl_2$] pH 6.8 unless otherwise noted. Concentration was determined after refolding by absorption at 280nm.

Circular Dichroism

Circular dichroism spectroscopy was performed using an AVIV model 215 CD spectrometer as previously described [92, 105]. Briefly, purified protein was centrifuged at 100,000xg for 30 minutes, then protein concentration was adjusted to approximately 0.3 mg/ml, and spectra between 260-190 nm obtained using a 1 mm quartz cuvette. Data was averaged from a minimum of three independent scans.

in vitro enzyme assays

Assays of purified enzyme activity were carried out in activity assay buffer [100mM NaOAc, pH 6.8, 0.1% Brij-35, 5mM $CaCl_2$] with the fluorogenic substrate (Abz-RVKRGLA-Tyr[3-NO_2]) as previously described [92, 105], except where use of the alternate substrate (Abz-RVARGLA-Tyr[3-NO_2]) is noted. Fluorescence was read using a SpectraMax –M2 spectrofluorometer equipped with a 96-well plate reader (Ex 320/Em 425). Data were averaged from a minimum of three independent experiments, and analyzed using GraphPad prism

To determine Km values, varying concentrations of the fluorogenic substrate was incubated with 0.02 EU [1 EU of enzyme produces 267 RFU/min] in activity assay buffer. IC_{50} values were determined similarly, with the addition of serially diluted amounts of either refolded WT or mutant PRO^{PC1} and a constant 15 uM fluorogenic substrate mixed in activity assay buffer, and mature PC1/3 added to initiate the reaction. Activity was assayed and analyzed as described above.

Discussion

Here we describe the biochemical effects of several single nucleotide polymorphisms in the propeptide of the neuroendocrine proprotein convertase PC1/3. Our preliminary results highlight the importance of the propeptide in overall PC function, as well as illustrating a phenomenon that has to date only been described in bacterial subtilases. While the propeptide domain is excised from the mature catalytic domain, our data none-the-less suggests that it plays a critical and multifaceted role in ensuring the correct function of this protease.

The first SNP we have considered is the $P_{76}S$, which is at the P5 position relative to the secondary cleavage site, as well as immediately C-terminal to a histidine residue proposed to attenuate the pH sensitivity of the conserved pH-sensing histidine [105]. The implications of the $P_{76}S$ polymorphism are two-fold; most basically, a nonpolar, conformationally rigid amino acid is substituted by a small polar amino acid, thus it is reasonable to suspect that the flexibility in the region would increase, and the presence of a polar, hydrophilic group may alter

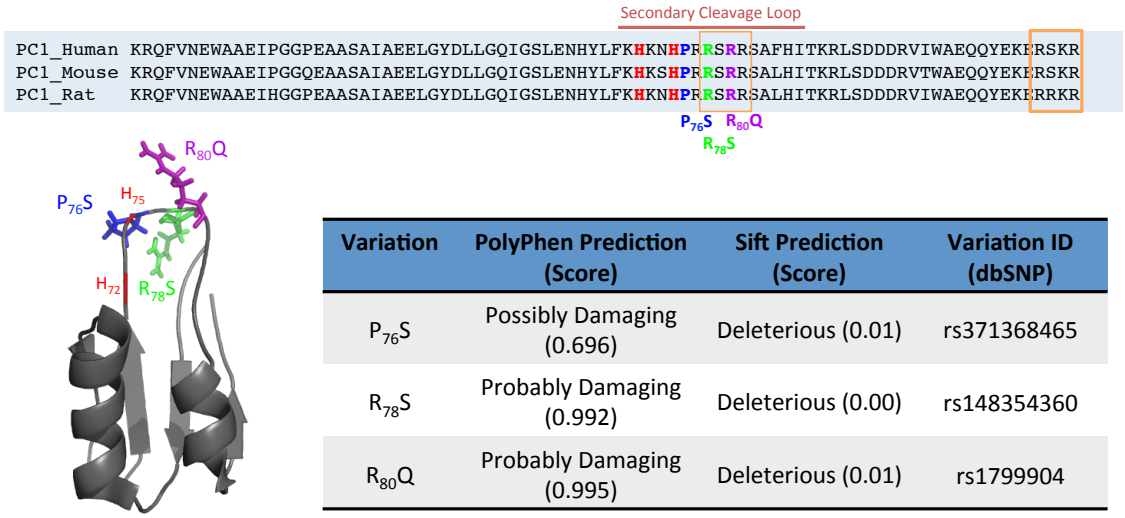
packing or further enhance conformational flexibility. Secondly, the juxtaposition of a histidine immediately preceding the proline has been shown to have implications for peptidyl-prolyl cis/trans isomerization [214]. Several proteins containing loop regions critical for regulation of protein function have proline switches near the base of the loop. When the proline is in the cis conformation, the loop is “kinked” and makes more contact with the body of the protein, however when the proline adopts the trans conformation, the loop extends from the body and is more mobile [232]. Peptidyl-prolyl isomerization is normally a slow conformational interconversion, however in peptides and proteins where a histidine immediately precedes the proline of interest, rates of interconversion increase 2-10 fold at pH below the pKa of the imidazole side chain, despite the general pH-independence of peptidyl-prolyl isomerization [212]. Interestingly, while histidine is a less commonly occurring amino acid, only occurring at a frequency of ~2.1% of all amino acids, histidine appears at a higher than expected frequency in the position immediately preceding proline [214]. Therefore it is tempting to speculate that the frequency of the His-Pro motif and its unique dynamic behavior may be an evolutionary adaptation to the need for regulation. As a final note, the three-dimensional fold of protein is important in determining conformation of a prolyl bond; it has been observed that the appointed peptidyl bond takes a cis conformation even when the original imino acid is substituted by another amino acid [233-235]. As amino acids most commonly take a trans conformation, being placed in cis may be sterically unfavorable.

The second potential implication we can hypothesize from these data is that alterations in the PC1/3 propeptide may affect propeptide processing, consistent with earlier reports, and thus may partially explain the changes in IC_{50} values we see (Figure 5.3)[218]. Mutation of Arg_{80} to Ala does not affect inhibition significantly [218], while mutation of this same position to Gln drastically decreases it, suggesting that there may be direct correlation between either size or charge of residues in the propeptide and active site that are critical to binding and inhibition. Accordingly, work on subtilisin indicates that mutant propeptides are able to recognize and more potently inhibit their cognate catalytic domains as compared to a catalytic domain produced with a wild type propeptide.

This observation portrays a third implication of propeptide polymorphisms that involves the role of the propeptide as a foldase. The propeptide imprints steric information upon the catalytic domain, a phenomenon termed protein memory. In this case, an identical sequence can give rise to altered three-dimensional structures dependent on whether its folding has been chaperoned by the wild type propeptide, or one that had been genetically altered by a single nucleotide polymorphism. Given the conservation of propeptide function between the bacterial subtilisins and PCs, it is reasonable to hypothesize that a similar effect of the polymorphisms could occur, thus imparting changes in structure, stability, and substrate specificity of the protease domain. Additionally, concomitant with the evolution of the secretory pathway in eukaryotic cells, we have proposed that the propeptides of the PCs have evolved to encode pH sensors to regulate their spatiotemporal activation. As alterations in the local

environment of the pH sensors could alter their pKas, a final possibility that must be considered is the effect these polymorphisms may have on pH sensing, and thus compartment-specific activation of the protease.

Ultimately, it is the effect on downstream substrate processing that is the true indicator of the biological relevance of these polymorphisms. While there are clear differences in processing of synthetic substrates, the effect on processing of natural substrates, for example proinsulin or POMC remain to be determined, and will be the true test of whether propeptide variations represent an interesting avenue for future research. At this point, many of these points are speculative; much remains to be done in order to more thoroughly characterize and understand the impact of these polymorphisms on the overall function of PC1/3. Specifically, it remains to be seen whether changes in the propeptide alter the pH-dependent unfolding of the isolated propeptides, and how these correlate with the *in vivo* pH-dependent activation profile. Furthermore, a close characterization of substrate affinity, and site-specific cleavage profiles of the protease chaperoned by the variant propeptide will undoubtedly reveal whether the SNPs have potential to cause alterations in prohormone processing that could drive endocrinopathies in human patients. Therefore, future work will primarily focus on *in vitro* studies, complemented by additional biophysical studies.



Variation	PolyPhen Prediction (Score)	Sift Prediction (Score)	Variation ID (dbSNP)
$P_{76}S$	Possibly Damaging (0.696)	Deleterious (0.01)	rs371368465
$R_{78}S$	Probably Damaging (0.992)	Deleterious (0.00)	rs148354360
$R_{80}Q$	Probably Damaging (0.995)	Deleterious (0.01)	rs1799904

Figure 5.1: Single Nucleotide Polymorphisms identified in the propeptide of PC1/3

A) Aligned sequences of the PC1/3 propeptide from human, mouse, and rat. The locations of the three SNPs identified are indicated below the sequences., and the two histidine residues responsible for pH sensing highlighted in red. Consensus cleavage sites are indicated by orange boxes, and the secondary cleavage loop indicated by the red bar above the sequences. **B)** Structure of the PC1/3 propeptide (1KN6) with locations of SNPs indicated. The pH-sensing histidines are indicated in red. **C)** PolyPhen and Sift scores for SNPs identified from dbSNP database.

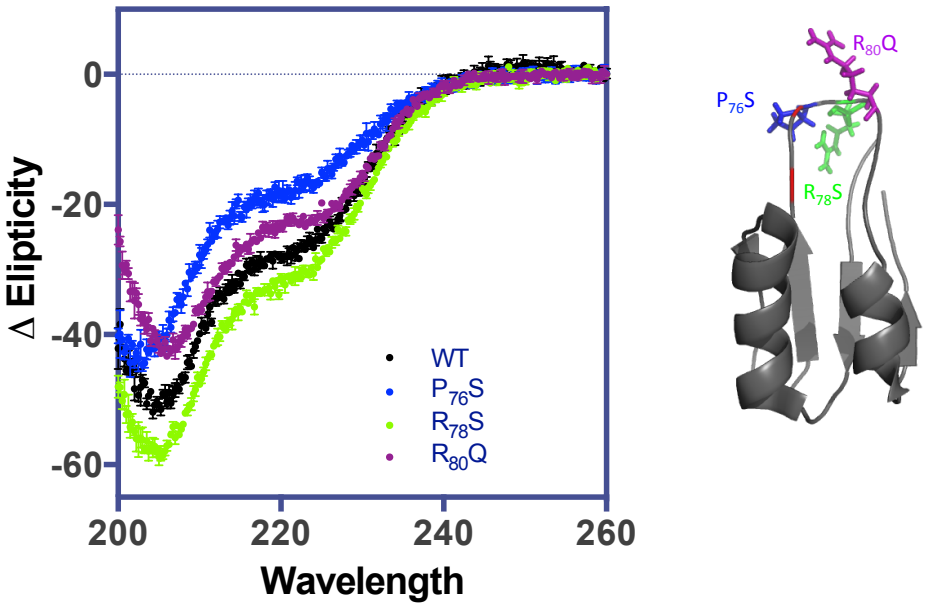
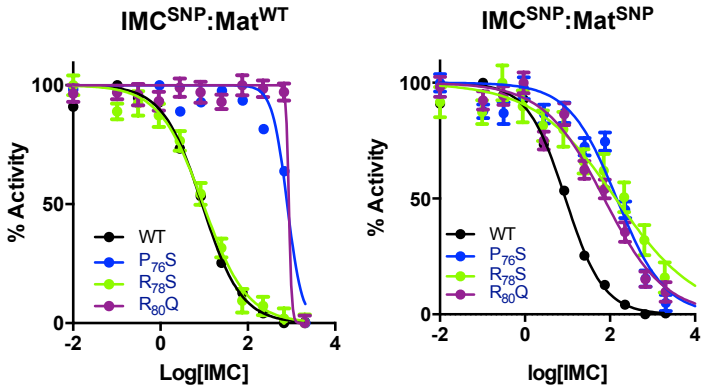
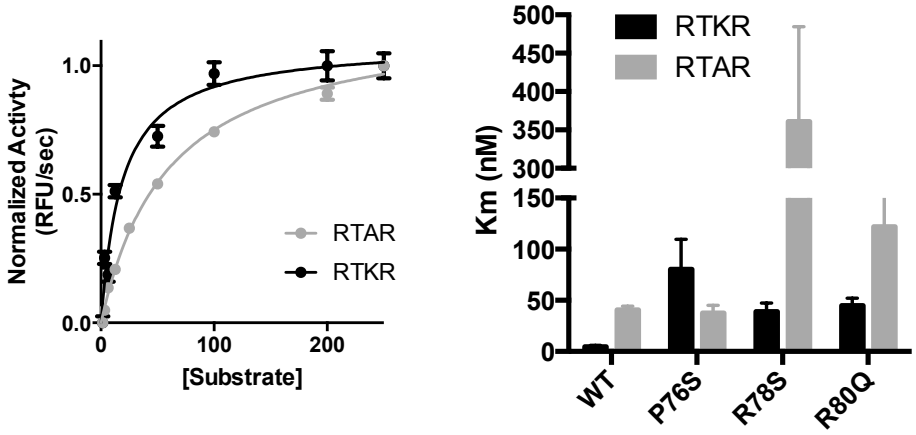


Figure 5.2: Secondary structure of isolated propeptides
Far UV (260 – 190 nm) circular dichroism spectra of the isolated wild type propeptide (black) or propeptides containing the indicated SNP variations; blue, $P_{76}S$; green, $R_{78}S$; purple $R_{80}Q$. The structure of the PC1/3 propeptide is provided for reference.



Propeptide	IC50 (nm)	
	Mat^{WT}	Mat^{SNP}
WT	8.94	
P76S	624.6	143.5
R78S	9.50	136.3
R80Q	872.4	72.38

Figure 5.3: IC_{50} values of propeptides for the catalytic domain of PC1/3
IC_{50} values were determined for the wild type propeptide (black) or SNP variants (blue, $P_{76}S$; green, $R_{78}S$; purple $R_{80}Q$) toward the protease folded by (right) the wild type propeptide (Mat^{WT}) or (left) their cognate propeptide (Mat^{SNP}). Mean values from three independent experiments performed in triplicate are given in the table.



Protease	RTKR (nm)	RTAR (nm)	Ratio (RTAR/RTKR)
WT	4.72	40.65	8.61
SNP2	80.31	37.65	0.47
SNP3	38.93	361.0	9.27
SNP4	44.89	122.0	2.72

Figure 5.4: Kinetics of the protease domain
Left: Representative enzyme progress curves for the wild type-folded PC1/3 for the two substrates. Right: Km values for each substrate; values are plotted as mean ± SEM and are the result of three independent experiments performed in triplicate.

Chapter VI: Proproteins, protons and pH sensing: A perspective on PC-related health and disease

Controlled proteolysis of secreted hormones, receptors, and enzymes is essential for optimal cellular function. Processing of these precursor molecules via endoproteolytic cleavage results in biologically active proteins, and has allowed evolution of elegant mechanisms through which precise spatiotemporal regulation of proprotein processing is ensured. Analysis of the human and mouse genomes reveals the presence of at least 990 and 791 distinct proteases and 198 and 193 non-peptidase homologues, respectively [236]. Approximately one third of these are serine proteases [237], which can be divided into two major families, those belonging to the trypsin or chymotrypsin fold, and those belonging to the subtilisin-like fold. A subset of the subtilisin-like serine proteases, the Proprotein Convertases (PCs) are involved in activating proteins involved in many aspects of cell biology and communication.

The now fundamental paradigm that active peptide hormones are derived from larger, inactive precursors was first proposed in 1967 following from two independent discoveries. While Don Steiner elucidated the processing of proinsulin to its active form, Michel Chrétien established that γ-lipotrophin (γ-LPH) and β-endorphin (β-END) were actually derived from the larger β-lipotrophin (β-LPH), and later that β-LPH as well as several other critical regulators of growth and metabolism, was itself part of an even larger precursor, proopiomelanocortin (POMC) [1, 238, 239]. It was not until nearly 25 years later that the first member of the proprotein convertase family was identified. Often called the canonical member of the PCs, furin was identified based on its homology to the yeast subtilisin-like endoprotease, kexin [240]. This discovery

spurred the search for other proprotein convertases, and over the past nearly 30 years, furin and its paralogues have been shown to play critical roles in maintaining homeostasis, as well as to be drivers of a diverse range of pathophysiological states, cancer progression and metastasis [241-244], obesity and diabetes [197, 228], hypercholesterolemia [245-249], hypertension [250-252], neurological function[116, 253, 254], as well as infection by viruses[88, 255-257], parasites[258], and bacteria[259-261]. Given their central role in biology, PCs are therefore prime targets for therapeutic intervention in a wide variety of human diseases.

The Typical Proprotein Convertases

The first seven PCs identified were furin, proprotein convertase 1/3 (PC1/3), PC2, PC4, PC5/6, PC7/8, and paired basic amino acid converting enzyme 4 (PACE4), and have been denoted as the Typical Proprotein Convertases (TPCs). They form a group of structurally conserved subtilisin/kexin-like serine proteases that function primarily within the secretory pathway, endosomes, and cell surface (Table 1). The substrates of the TCPs can be largely identified by their consensus dibasic cleavage motif, (R/K)-XX-(R/K). PC1/3 and PC2 are the convertases responsible for the processing of polypeptide hormones and neuropeptides, including insulin and POMC, and are found in the regulated secretory pathway of endocrine and neuroendocrine cells [262]. Four PCs, furin, PC5/6, PACE4 and PC7 are near-ubiquitously expressed and are responsible for the processing of receptors, ligands, enzymes, and

growth factors that take place in the TGN, endosomes and cells surface of the constitutive secretory pathway [13]. SKI-1 and PCSK9, the remaining two members of the PC family, are members of the pyrolysin and proteinase K-like clades of subtilases, respectively, and are unique in both their preferred cleavage motif and proteolytic activities. SKI-1 cleaves client proteins after non-basic residues, while PCSK9 to our current knowledge does not cleave any substrates other than itself. Given that he structure of PCSK9 is similar to furin, the loss of the ability of PCSK9 to cleave other substrates *in trans* demonstrates an interesting aspect of evolution, wherein the biological function of the protease scaffold has instead evolved to a ligand that binds to the low-density lipoprotein receptor to induce its internalization and degradation [16]. An analysis of the MEROPS database established that approximately 20% of proteins that show sequence similarity with proteases are classified as non-peptidase homologues. Hence, while no less interesting than the TCPs, SKI1 and PCSK9 are more closely related to the plant and bacterial subtilases, and as such will not be discussed here in detail.

The TPCs are a ubiquitous and critically important family of proteases within eukaryotes. Like their substrates, the proprotein convertases must also undergo multiple proteolytic cleavages, most of which involve their N-terminal propeptides, before the protease domain becomes fully active [122]. Following translocation into the ER, the zymogens are folded with the guidance of the propeptide, which acts as an Intramolecular chaperone (IMC). Correct folding of the catalytic domain allows the first autoproteolytic cleavage; while the

propeptide and protease domain are no longer covalently linked, the propeptide remains associated to the protease, maintaining it in an inactive state as it exits the endoplasmic reticulum. Subsequent secondary cleavage of the propeptide is required to release the fully mature, active enzyme, and occurs in a compartment- and pH-specific manner. Since premature protease activation can lead to inappropriate downstream protein activation, sorting, or degradation, it should come as no surprise that mutation or deficiencies of specific PCs have diverse and often devastating effects on organismal homeostasis. The role of PCs in health and disease has been thoroughly reviewed previously [115, 161, 183, 263], and new reports continue to be published almost weekly. A large amount of attention in recent years has been paid to the role of furin in promoting tumor growth and metastasis, owing to its role in cleaving a diverse range of growth factors, receptors, scaffolding proteins and others. TPCs have also been implicated in various neurologic diseases, including Alzheimer's [264-266] and recently Huntington's disease [267].

Proprotein Convertases in mice and men

In vitro experiments have demonstrated that the TPCs possess closely related and even redundant biochemical properties, and often share substrate molecules. However, the result of genetic knockout of individual TPCs in mice argues for substrate specificity [162]. Further insight into the unique role of the TPCs is offered by the phenotypes of individuals who have either mutated or

missing PCs; interestingly, in many cases the mouse models do not phenocopy the human (Table 2).

Furin, PC5/6 and PACE4 are essential for normal mammalian development, and as such knockout in mice have severe consequences. Furin deficient mice die at embryonic day 10.5 due to failure of chorioallantoic fusion, ventral closure and axial rotation [193], while mice lacking PC5 die at birth due to multiple craniofacial and pattering abnormalities [268, 269]. No humans deficient in furin have been identified to date, a fact that reflects its central role in development. However, expression of furin has been reported to be misregulated in several human diseases, where upregulated furin promotes cancer metastasis [270], formation of atherosclerotic lesions [271], and is hijacked by viral, bacterial and parasitic infections for processing of virulence factors [272]. Furthermore, as furin is directly involved in the regulation of the renin-angiotensin system, several single nucleotide polymorphisms in the *furin* gene have been associated with elevations in both systolic and diastolic blood pressure [250, 251, 273]. Furin also moonlights as a chaperone that facilitates folding of MMP28, independent of its catalytic activity [79]. In this case, furin deficiency removes a structural scaffold that affects the function of a protein essential to a seemingly unrelated biological pathway, paralleling the role of PCSK9 in lipid metabolism. PC5/6 is likewise ubiquitously expressed, but differs in trafficking based on alternate splice variants; the shorter PC5/6a is secreted, while the longer PC5/6b is membrane anchored. Like furin, due to its critical role in development, no patients lacking PC5/6 have been identified; however several polymorphisms have been

associated with pathologies, many with developmental and neurologic implications. PC5/6 regulates *hox* gene paralogues via growth differentiation factor-11, and has been implicated in aberrant anterioposterior patterning causing anorectal atresia [269]. Recent meta-analyses have linked SNPs in intronic regions with onset of Alzheimer's disease and amyotrophic lateral sclerosis [274, 275]. Knockout of PACE4 is less severe, however mice still exhibit complex craniofacial malformations, and left-right patterning defects including heterotaxia, cyclopism, and pulmonary isomerism [276]. Similarly, human polymorphisms have been associated with antero-posterior and left-right axes specification [277], increased expression of PACE4 has been found in ovarian cancer and squamous cell carcinoma [278, 279]. Most interestingly, a polymorphism in PACE4 causing variable splicing of the coding region may alter the trafficking of the protein from primarily intracellular to secreted, and contribute to the pathogenesis of osteoarthritis by promoting aggrecan breakdown [280-282].

PC1/3 and PC2 are closely related, and often perform opposing but complementary functions to maintain homeostatic balance via processing of multiple hormone precursors, including proinsulin, proopiomelanocortin, prorenin, proenkephalin, prosomatostatin, progastrin, proglucagon and proghrelin [219, 283, 284]. Accordingly, mutations and polymorphisms in both proteases have been shown to cause complex and often severe metabolic derangements, adrenal hyperplasia [285], gastrointestinal carcinoids [286], pituitary adenomas [287], and cancers [288, 289] in humans. To date, four patients have been

identified with mutations in PCSK1 resulting PC1/3 deficiency, causing early-onset obesity, hyperphagia, hypoadrenalism, reactive hypoglycemia, malabsorbtive diarrhea, hypogonadism, and diabetes insipidus [199, 201, 202, 224]. A number of single nucleotide polymorphisms in PCSK1 have likewise been identified and described; these polymorphisms affect catalytic activity and substrate processing, and have been associated with increased risk of diabetes and obesity [126, 290-292], as well as premature ovarian failure [225] and coronary artery disease [226]. In many cases, even minor alterations in PC1/3 activity can shift the homeostatic balance of an organism, predisposing them to metabolic derangements [126, 224]; as such, PCSK1 has been called the third most important gene contributing to extreme obesity [197]. Recent evidence suggests PC1/3 and PC2 may also be active in the innate immune system, playing a role in the generation of the inflammatory response via regulation of cytokine release from macrophages [293]. The PC1/3 knockout mouse is growth retarded, and has deficits in insulin synthesis and limited POMC processing, however interestingly, maintains normal glucose homeostasis and is not obese [194, 227, 228, 294], observations that are strikingly different from patients with PC1/3 deficiency. Similarly, the PC2 knockout mouse exhibits restricted growth, fasting hypoglycemia, and multiple substrate processing defects [295-298]. While both single knockouts are viable, the compound PC1/3/PC2 null mouse is lethal, supporting the conclusion that lack of one can be partially compensated for by the other, but not by other convertases [116].

Less is known about the last two TPCs, PC4 and PC7. PC4 expression is restricted to germ cells, and few natural substrates have been identified to date [299-301], but PC4 null mice are infertile [302]. PC7 is ubiquitously expressed, and appears to be redundant in function to many of the other PCs; however, a few unique substrates have been identified recently. Notably, PC7 may have a role in iron metabolism via release of soluble haemojuvelin and soluble transferrin receptor from the cell [303-305]. Additionally, silencing of PC7 expression in mice results in an anxiolytic and novelty-seeking phenotype, suggesting a role for it in regulation of dopaminergic circuits [306]. These data support the hypothesis that despite their *in vitro* overlapping specificity, there are mechanisms *in vivo* that delineate substrate preference.

Signaling through changes in proton concentration

The dynamic intraorganellar localization of the TCPs is likely a critical factor in determining the *in vivo* processing individualities of each protease. As mentioned above, the TPCs are activated in a pH-dependent and compartment specific manner in a series of autocatalytic processing events that are driven by its propeptide. While not a part of the mature, functional protease, the propeptide encodes the necessary information required to dictate pH-dependence of activation; it has been demonstrated that swapping the propeptides of PC1/3 and furin is sufficient to reassign the pH-dependence of activation *in vitro* and *in vivo* [92]. Following folding of the protease, the first autoproteolytic cleavage breaks

the peptide bond joining the propeptide to the protease, however the propeptide remains associated with the mature protease, occluding the active site until the complex is properly trafficked to the appropriate organelle. The second autoprocessing event is triggered by the protonation of a conserved pH-sensing histidine within the propeptide [103, 105] that drives local structural changes that expose the cleavage site [104]. It has been hypothesized that the propeptides of the TCPs have been enriched in histidine residues, relative to their catalytic domains, as a method of encoding and fine-tuning this pH sensitivity [95] such that despite their conserved pH sensor, each PC maintains a distinct organellear domain in which it is active [105, 106].

In mammalian cells, the nucleus, endoplasmic reticulum and cytoplasm have near-neutral pH; mitochondria are more basic, while the Golgi network, lysosomes and endosomes are progressively more acidic [97]. Maintenance of appropriate pH within individual membrane-enclosed compartments is critical for the normal physiology of the cell and organelles. For example, when the pH gradient across the Golgi is collapsed, post-translational modifications and processing of secreting is impaired, cargo is misdirected, and the integrity of the organelle itself is compromised [206]. Many of the tools required to fully understand the pH changes that happen in the physiologic and pathologic state are still in their infancy and the specifics of these changes in individual disease have been reported on elsewhere [182, 206]. Intracellular pH is tightly regulated to maintain homeostasis via H+ ion transporters (for example, the vATP-ases) in conjunction with various ion channels and transporters[75, 307]; as protons

function as second messengers in many signaling pathways, perturbation of intracellular pH can drastically affect cellular functions. Functionally, changes in pH are required for progression through the cell cycle, differentiation, migration and chemotaxis, proper function of receptor ligand pairs, rates of protein synthesis and apoptosis. Additionally, acidification of the cytosol can disrupt membrane traffic along the endocytic and biosynthetic pathways as a result of defective assembly of clathrin coats at the plasma membrane, and redistribution of endocytic compartments within certain cell types [308-311]. Furthermore, dysregulated pH homeostasis is a hallmark of several disease states, including cancer, neurodegenerative disorders, Dent's disease, and cystic fibrosis [109, 182, 312-314].

Alterations in steady-state organellear pH have likewise been suggested to contribute to the pathology of several diseases, however specific examples remain elusive [75, 307]. While there has been no definitive link between alterations in cytoplasmic pH (pHi) and alterations within organellear compartments, experiments in yeast suggest that vacuolar pH can be perturbed by acidification or alkalinization of the medium, with a particularly notable effect seen upon alkali stress, where median vacuolar pH increased from 5.27 to 5.83 [315]. Several other observations likewise support this hypothesis; Inositol 3,4,5,6-tetrakisphosphate inhibits acidification of insulin secretory granules by blocking ClC channels [316]. Reduced acidification in turn reduces exocytic release of insulin from granules, thus may be a driver of the etiology of type II diabetes mellitus. Levels of Ins(3,4,5,6)P4 are elevated in response to a

decrease in ATP/ADP ratio[317], which is known to be altered in a variety of disease states including cancer, cardiovascular disease, aging, and diabetes itself [318]. Similarly, activation of protein kinase C, which is likewise frequently dysregulated in cancer, diabetes, cardiovascular disease and many other pathologies [319] has been noted to depress Golgi acidification [320]. As the pH of the secretory pathway is under regulation of multiple second messengers, and given the multiplicity of steps controlled by luminal concentration of proteins, it is reasonable to suspect imbalances in these secondary messengers could result in alterations in biosynthesis, processing, and export of secretory products, causing multiple distinct yet comorbid diseases, even if a direct link has not yet been demonstrated.

An example of altered pH homeostasis that has received a great deal of attention in the last decade has been the role of intracellular pH on insulin secretion, and how this may contribute to our understanding of diabetes mellitus. Studies have shown that a decrease in intracellular pH is favorable for glucose-stimulated insulin secretion, while intracellular alkalization is inhibitory [321-323]. More recent data has demonstrated that intracellular pH determines the ability of nutrients to stimulate insulin secretion; the optimal pH range for insulin secretion is acidic, and even a slight increase of pH out of the physiologic range is enough to strongly inhibit nutrient-stimulated insulin secretion [324]. Therefore, it is possible that a small defect in intracellular pH regulation could cause a significant defect in insulin secretion. This is particularly relevant in cases of non-insulin dependent diabetes mellitus (NIDDM), where islet pH can decrease below the

normal range due to plasma acidosis, or increase either due to overcompensation for acidosis or due to hyperglycemia. Such misregulation undoubtedly contributes to the secretory defects of NIDDM directly, but may also explain the aberrant and incomplete processing of insulin seen in these cases due to inappropriate activation of the proprotein convertases responsible for this conversion.

The question of the effect of altered pH homeostasis on the activation of the PCs is an important point to consider. It has long been known that many prohormones as well as their processing enzymes aggregate in acidic conditions; a consistent finding in vitro has been that pH perturbation diverts hormone precursors into the constitutive secretory pathway, thus preventing their proteolytic maturation via mistrafficking [325]. Whether through altered trafficking or a failure to maintain appropriate pH gradients, or the synergistic effect of both, the premature exposure of the PC inhibition complex to an acidic environment could cause early activation, while neutralization could prevent their activation altogether. In an earlier piece of work, we noted that there appears to be a window of activation for furin, suggesting that in addition to preventing furin from being active before it reaches the TGN, the propeptide can likewise ensure that it remains inactive should it exit this compartment without being properly activated for any reason [104].

Evolution of environmental sensing

Given the importance of pH regulation throughout the cell and across the secretory pathway, it is not surprising that many proteins have evolved ways of exploiting the changing pH for their own function and regulation. Amino acid encoded pH sensors have been shown to play a role in regulation of a wide variety of proteins, modulating structure, stability, solubility, and binding in a protonation-state dependent manner. The pKa values of ionizable residues in folded proteins can be strongly influenced by the local environment. The three major factors affecting these values are charge-charge interactions, charge-dipole interactions, and the Born effect [326, 327]. Furthermore, changing environmental pH can modify the protonation state of a residue, causing a conformational change that then alters the environment of the residue, changing its pKa value [328]. Tanford has shown that if the equilibrium between two conformational states of a protein is pH dependent, then at least one titratable group must have a different pKa in the two conformations [329].

A major advantage of protonation as a post-translational modification is the potential for rapid temporal responses. Protons can diffuse through water quickly (via Grotthuss diffusion), possibly facilitated by short-range motion in proteins. Protons also present an elegant means of regulating protein activity and interactions. A proton is an exceptionally small single subatomic particle that can result in a reversible chemical change with significant effect on electrostatics, driving changes in protein structure, dynamics and interactions. Additionally, in contrast to other post-translational modifications, addition and removal of a

proton does not require an enzyme. Signaling specificity is achieved by only a minority of sites in proteins titrating within the physiologic pH range, thus the variation in pH across subcellular organelles makes spatial regulation possible. Used in multiple signaling modes, the Bohr effect in hemoglobin is the classic example of a single-site, coupling of protonation and conformational equilibrium, where a His-Asp bridge is formed with protonation of the histidine at decreased pH to facilitate oxygen off-loading[330]. Additional examples include regulation of the catalytic activity and substrate binding of glycinamide ribonucleotide transformylase, protein-protein interactions of the neonatal Fc receptor, and aggregation of the mammalian prion protein PrP^{C} [182, 207].

pH-sensors are however a double-edged sword; while an elegant evolutionary solution to the problem of spatiotemporal regulation, their proper functioning is inextricably linked to the pH homeostasis of the cell. Whether this is a ‘chicken and egg’ situation or if there is a yet-undescribed unifying theory remains an active area of questioning, but given current evidence for the role of pH dynamics in regulation of such a broad range of cellular processes, it is reasonable to hypothesize that dysregulated pH drives dysfunction of proteins that leads to further perturbations of pH homeostasis in a feed-forward fashion.

It is interesting to note that the subtilase fold does not mandate the requirement of a propeptide; several bacterial subtilisin-like proteases fold to their active conformations without the assistance of dedicated propeptide domains. Thus, from an evolutionary standpoint, eukaryotic PCs could have evolved from either propeptide-dependent or propeptide-independent subtilases.

Integrating chaperoning and pH sensing

As mentioned above, the propeptides of the PCs act as intramolecular chaperones (IMCs), and function as post-translational modulators of protein structure and function. In contrast to canonical molecular chaperones, which chaperone folding *in trans* and are not usually substrate specific, intramolecular chaperones are encoded within the primary sequence of a protein as an N- or C-terminal sequence extension. While not a direct contributor to the proteins function, they are nonetheless essential for folding of the functional protein *in cis,* and function only once to fold the cognate protein before they are degraded. It is interesting to note that mutations in the IMC can cause misfolding of the functional domain, resulting in distortion of their function, even if the primary structures of functional domains are identical. Such mutations are termed "protein-memory mutations". This phenomenon of imprinting is imparted in the late stages of folding, and was first described by Inouye and colleagues in studies of subtilisin; an identical peptide chain can fold into multiple altered, but active, conformations under physiological conditions through point mutations in the propeptide (Figure 6.1) [48, 112]. Not only do these folded proteins differ in activity, they are able to differentiate between wild type and mutant propeptides and specifically bind to the propeptide that mediated their folding with higher affinity. The altered protein also displays specificity that is different from the wild type protein; in the case of the PCs, this could lead to incorrect hormone processing, driving hormonal disorders.

Numerous single nucleotide polymorphisms (SNPs) have been identified in the PC family, and several characterized as associated with various human diseases, as noted previously (Table 2). The most thoroughly characterized SNPs have been reported in PC1/3, and have been associated with obesity and diabetes in multiple genome-wide association studies [126, 290-292]. Obesity is a common disorder affecting more than 35% of American adults [331], and involving multiple genetic factors. The role of PC1/3 in regulation of appetite, and consequently obesity, via its action on key peptides in the leptin-melanocortin pathway has been well studied [332]. Several patients with nonsense or missense mutations yielding non-functional PC1 have been characterized before, however, the role of SNPs is less clear; there have been a veritable wealth of studies across ethnic groups attempting to ascertain the association between certain PCSK1 variants. Original studies in European populations indicated that common variants (rs6232 and rs6234-rs6235) contributed to obesity risk [126], however studies replicated in other European, Asian, and Mexican populations have had mixed results with varying SNPs being more or less correlated with obesity risk [333]. While interpretation of these results is obscured by differences in study methodologies and unknown heterogeneity in genetics, lifestyle and environmental factors, this highlights the fact that SNPs within *PCSK1* contribute modestly to obesity within a multiethnic population such as the United States, and variably dependent on age, genotype, and unique ethnic background.

Of particular interest in our research is the role of the SNPs identified in the propeptides of the PCs. As the IMCs are requisite for folding, it is likely that

the altered IMCs are affecting structural heterogeneity despite the sequence heterogeneity of the proteases. Additionally, it would be reasonable to hypothesize that these SNPs alter the structure of the IMC itself, thereby altering the local environment of the histidine-encoded pH sensors, and altering the stability of the propeptide. These changes undoubtedly have significant effects on protease function and substrate specificity that could have marked consequences for cellular and organismal homeostasis. Again, PC1/3 has been paid the most attention in recent literature. Engineered single amino acid substitutions within the propeptide of PC1/3 have been shown to alter the inhibitory profile and mechanism of its cognate protease [218]. Furthermore, the most common polymorphism in the propeptide reported to date encodes an arginine to glutamine mutation, which was shown to alter the processing, trafficking and catalytic activity of the protease [113]. Expanding on these initial observations, we have recently undertaken further biochemical and biophysical characterization of two additional SNPs that, based on *in silico* predictions, may be detrimental to protease function (Williamson, manuscript in preparation). Our results demonstrate that the polymorphisms alter the structure of the isolated propeptides, with two critical results; first, the affinity of the variant propeptides for their cognate catalytic domains is significantly less than the wild type, thus their ability to act as an inhibitor of protease activity is likely altered, a point that may alter the spatiotemporal regulation of their activity. Secondly, the catalytic efficiency of *in vitro* substrate processing is diminished. This may not only impact *in vivo* substrate processing, but also the autocatalytic removal of the propeptide.

While it remains to be seen whether the altered structure of the propeptides impact the pKas of the pH sensing histidines, or alter the *in vivo* behavior of the catalytic domain, we believe the evidence strongly supports the biological relevance of these polymorphisms, and thus underscores the importance of building a greater understanding of the interplay between secreted proteases and their propeptides.

Proteins evolution though successive generations can occur through genetic drift and natural selection. Such changes, which result from point mutations, insertions or deletions, can alter proteins such that their common origin cannot be detected from sequence, even though the protein may retain its original structure and function. We propose that this is the case for the propeptides of the subtilase family; while there is little sequence homology between the prokaryotic subtilisins and the eukaryotic homologues, the overall structure of the domain has been conserved. Even in considering only the propeptides of the seven TPCs, we see much less sequence homology than would be expected.

It has previously been demonstrated that there is an increase in histidine content in the propeptides of eukaryotic proteases that are activated under acid conditions. This was first noted in the proprotein convertases, then confirmed in the unrelated cathepsin family, which is also activate at low pH [95]. These findings suggest that histidine enrichment in regulatory peptides is an example of convergent evolution. Presumably, histidine enrichment in propeptides is a facile way to encode pH-regulation; the propeptides of the subtilases are under weak

evolutionary pressure, with the only requirement being a general maintenance of structure, as compared to their catalytic domains, which must maintain the precise arrangement of the catalytic triad. Furthermore, the structure of propeptides is optimized to balance structural integrity with flexibility, in order to maintain tight inhibition while still allowing for activation by cleavage [66]. By selecting for a histidine at a position that can accommodate an uncharged imidazole ring, but not a positive charge, pH, protonation and activation are all immediately coupled to a structural change. Introduction of additional histidines and fine-tuning of the stability of the propeptide allows for facile yet precise modulation of the pH of activation. Furin is ubiquitously expressed at various levels in all tissues, thereby explaining its widespread role in the processing of various proteins; similarly the variety of substrates and their distribution throughout the endosomal recycling pathway and the cell surface indicated that furin's processing role took place at near-neutral pH. In contrast, processing of most hormone precursors occurs in the secretory granules, and is understood to occur at acidic pHs, suggesting that the cognate processing enzymes would likewise be active in these circumstances. Again, the observation that the cleavage of different substrates is delineated by compartmental and environmental factors suggested a conserved mechanism by which PC activity was proscribed spatiotemporally; the necessity of ensuring activation coincided with localization may have put an evolutionary constraint on the pH dependence of these proteins such that they matched the subcellular differences in pH [334].

Protein duplication allows for complexity within pathways and systems. As the duplicated copy is freed from the selective pressure to maintain broad functionality, it is able to accumulate mutations at a higher frequency, until it assumes a novel selectable function, or is silenced via mutational inactivation [335]. Proteins change through successive generations through random drift and natural selection, most frequently through point mutations, insertions and deletions that alter the sequence but not overall structure of the protein [336]. It is the balance between changing sequence and static structure that is of primary interest in the case of the propeptides of the subtilases, and especially the proprotein convertases. While the structure of the propeptide has been conserved between prokaryotes and eukaryotes, the differences in the sequence of the propeptide reflects the specialization of the PCs, specifically their substrate preference and pH-dependent activation. The pattern exemplified is an elegant one; the emergence of progressively higher complexity is built upon a foundation of simplicity.

Therapeutic potential

The prohormone theory opened the doors to fifty years of research of the proprotein convertases, a field that is still vibrant and active. A review of the literature reveals an impressive array of contributions to understanding the role of these proteases in human health and disease. Given the enormous advances gained from knowledge of the characteristics of the PCs, it is tempting to speculate about the power of emerging therapies to regulate their activities in

various pathologies. As discussed to a great extent here and elsewhere, the propeptide of the PCs plays an important role in their regulation. Clinical and mouse studies have revealed that the proprotein convertases are at the crossroads of several pathways implicated in disease, including cancer progression and metastasis, obesity, diabetes, hypertension and neurologic function, as well as viral, bacterial and parasitic infection. On this basis, inhibition of the PCs represents a viable therapeutic target.

Blocking catalytic activity of PCs using protein-based inhibitors has been established to reduce the progression of cancer, cardiovascular disease and viral infections, both *in vitro* and *in vivo* [188, 263]. Protein-based active-site directed inhibitors of PCs have been engineered by substituting the consensus substrate binding site of PCs into protease inhibitor scaffolds such as leech eglin-C, turkey ovomucoid, α_2-macroglobulin, and alpha-1-antitrypsin [337]. Studies have demonstrated that engineered alpha-1-antitrypsin, when delivered into human lymphocytic cell lines or primary human lymphocytes using recombinant simian virus-40-based vectors, can function as a potent inhibitor of PCs [338]. However, inhibitors that contain the consensus substrate binding sequence engineered into protein scaffolds are non-selective, and are recognized by all PCs which display conserved 3D structures, enzyme mechanisms, and substrate specificity.

Propeptides have been used *in vitro* and *in vivo* as protein-based inhibitors of furin. For example, adenovirus mediated hepatic over-expression of the furin propeptide in mice can significantly lower plasma LDL-c level and reduce atherosclerotic lesions [339], while injection of carcinoma cells over-

expressing the propeptide into nude mice delays tumor development and reduces liver metastases [188]. Hence, targeted delivery of the prodomain appears to be a therapeutically beneficial and feasible strategy for targeting PCs [191, 263, 340, 341].

While targeted inhibition of these enzymes may be acutely therapeutically beneficial, it is not without risk to organismal homeostasis; the propeptides of furin and PC1/3 display significant cross reactivity with PC-paralogues [188]. Structural analysis of inhibitory complexes establishes that propeptides interact tightly with catalytic domains at two distinct sites (Figure 6.2). The C-terminus of the propeptide interacts with the conserved substrate binding sites in the catalytic domain and overlaps with region where alpha-1-antitrypsin binds with PCs. However, propeptides form a second site of contact that is facilitated by beta sheets that interface with two alpha helices of the catalytic domain. Unlike the first site of contact with the substrate-binding region, amino acids that form the second site of contact are less conserved in both propeptides and cognate PCs [122].

Studies, including those from our group, have established that reducing interactions at either site dramatically diminishes binding affinity [50, 56, 342, 343]. Conversely, increasing interactions at the less conserved interface, by optimizing the propeptide sequence for steric complementarity with its cognate protease, helps differentiate between paralogues. It is also possible that point mutations distal to the sites of interaction between propeptides and PCs may alter binding affinity by affecting the interface through long-range interactions. It

is tempting to hypothesize that only a few amino acids in the propeptide domain are essential for binding, and by targeted mutations, we can increase binding affinity of propeptides for proteases, to make engineered propeptides dead-end inhibitors that exclusively target the catalytic domains of PC1/3 and furin, respectively. Thus it may be possible to exploit what we have learned about propeptide dependent folding and activation to create potent protein-based inhibitors that can target specific convertases. For example, proteases cleave at regions that are mostly unstructured or partially structured. Conversely stabilizing the hydrophobic core will enhance the proteolytic stability of the proteins. The propeptides represents a globular protein with a fairly tight packing of non-polar amino acids within a hydrophobic core. Studies have established that: (i) formation of this core is essential in folding, stabilization, and conformational specificity of propeptides in subtilases [48, 112]**;** (ii) intrinsic stabilities of the isolated propeptides of furin and PC1/3 differ, correlating with their optimum pH for activation [92]; and (iii) enhancing propeptide stability increases binding affinity [104]. We hypothesize that increasing the thermodynamic stability of propeptides will enhance resistance to proteolytic degradation and increase binding affinity. Why do we believe this strategy of structural stabilization will work? Our work has demonstrated that removing the pH sensor in furin stabilizes the propeptide and enhances its binding affinity towards the catalytic domain. Furthermore, preliminary informatics analysis shows that the propeptides of PCs and subtilisin are conserved in several protein families including molecular chaperones, intramolecular chaperones, metal binding proteins, RNA binding

proteins, membrane transporters, and enzymes. All these proteins have a conserved ferredoxin-like fold (Figure 6.3), but display different stabilities. Hence, analysing structurally conserved folds, manually or through Rosetta design [344], will allow for optimization of the hydrophobic cores in propeptides of PCs.

An alternate strategy for therapeutic manipulation, particularly applicable to the case of polymorphisms causing altered folding, trafficking and/or pH-dependent activation, is restoration of proper functioning via exogenous propeptides. Understanding pH-sensitive regulation at the structural level facilitates the rational design of therapeutics. Gain or loss of pH sensor function can drive pathologies, or be engineered to resist such pathologies. Much remains to be learned about the role of pH in regulation protein function and cellular processes, and of the link between the two. Additionally, as substrate specificity is governed by matching of the charge distribution pattern within the substrate binding pocket [10], and the propeptide is responsible for guiding the folding of the catalytic domain, a better understanding of how changes in the structure of the propeptide are mirrored in the fold of the protease may allow us to predict the deleterious effects of polymorphisms, as well as engineer exogenous propeptides to “reprogram” the specificity of a convertase.

In conclusion, much remains to be learned and discovered in the field of proprotein convertase biology, an endeavor that will undoubtedly be facilitated by an increased appreciation for the biochemical and biophysical structure-function relationship of this family, and of proteases as a whole. It is my hope that the body of work presented here will spur the investigation and development of

effective therapeutics that can be taken from the bench to the bedside in the future.

Table 1: Tissue distribution, subcellular location, and substrate of the Proprotein Convertases

Proprotein Convertase	Subcellular Activation Site, (pH)	Tissue Distribution	Target substrates
PC1/3	Secretory granules (~ 5.5)	Neural, Endocrine	Prohormones
PC2	Secretory granules (~5.5)	Neural, Endocrine	Prohormones
Furin	TGN, cell surface, endosomes (~ 6.5)	Ubiquitous	Growth factors, receptors, adhesion molecules, metalloproteinase, vATPases, viral glycoproteins, bacterial toxins
PC4	Cell surface?	Germ cells	IGF2, PACAP, proteins involved in sperm motility, reproduction
PC5/6	Cell surface, ECM (~6.5)	Adrenal cortex, intestine, kidney, ovary	Growth factors, receptors, adhesion molecules
PACE4	ECM, cell surface, (~7.5)	Muscle, heart, pituitary, intestine, cerebellum, kidney	Growth factors, metalloproteinases, viral glycoproteins
PC7	TGN, cell surface, endosomes (~ 6.5)	Ubiquitous	Receptors
SKI-1	Cis and medial-Golgi (~ 6.5)	Ubiquitous	Membrane-bound transcription factors?
PCSK9	TGN, extracellular (no proteolytic activity)	Liver, intestine, kidney, CNS	Binds LDL receptor, VLDL- receptor

Table 2: Consequences of mutations in the Proprotein Convertases

Proprotein Convertase	Mouse Phenotype	Clinical Phenotype	
		Deficiency	Mutation/SNP
PC1/3	Knockout: Dwarfism, increased pre- and perinatal mortality, hyperproinsulinemia and multiple hormone processing defects [194, 227] $N_{222}D$: Obese	Early onset obesity, hyperphagia, impaired glucose homeostasis, hypotonatodropic hypogonadism, hypoadrenalism, malabsorbtive diarrhea [197, 201-203, 217, 224, 345-347]	Obesity-related traits [126, 195, 199, 333, 348-350], Diabetes-related traits [291, 292, 351], Age at natural menopause [245], Premature ovarian failure [225], Cardiovascular risk factors [226]
PC2	Knockout: chronic hypoglycemia, growth retardation, decreased adiposity, multiple processing defects [296-298, 352-357]	Not described	Diabetes related traits [358-361], kidney function [362, 363], cardiovascular risk factors [364, 365]
Furin	Knockout: Embryonic lethal; failed axial rotation, ventral closure and chorioallantoic fusion, cardiovascular defects [193]	Not described	Blood pressure [251], hypertension [250], HPV infection [366]
PC4	Decreased fertility [302]		Intrauterine growth restriction [367]
PC5/6	Knockout: Increased prenatal mortality due to craniofacial and patterning defects [268, 269],		Possible association with caudal regression syndromes, patterning defects [269], ALS onset [275], Alzheimer's disease [274], HDL levels [368]
PACE4	Knockout: Embryonic defect anterioposterior and left-right patterning, complex craniofacial and CNS defects [276]	Not described	Osteoarthritis [280, 281], hypertension [369]
PC7/8	Anxiolytic and novelty-seeking phenotype [306]	Not described	Iron homeostasis [303], cardiovascular risk factors [370]

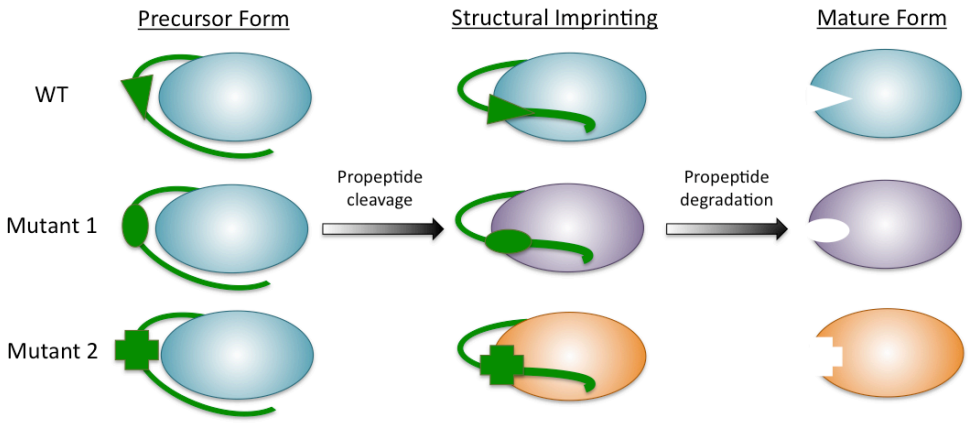


Figure 6.1: Protein Memory
Schematic of the results of mutations within a propeptide that acts as an IMC on its cognate catalytic domain. Propeptides are represented as green lines; Subsequent to chaperoning folding, the propeptides are removed, however structural information is imprinted upon the mature catalytic domain.

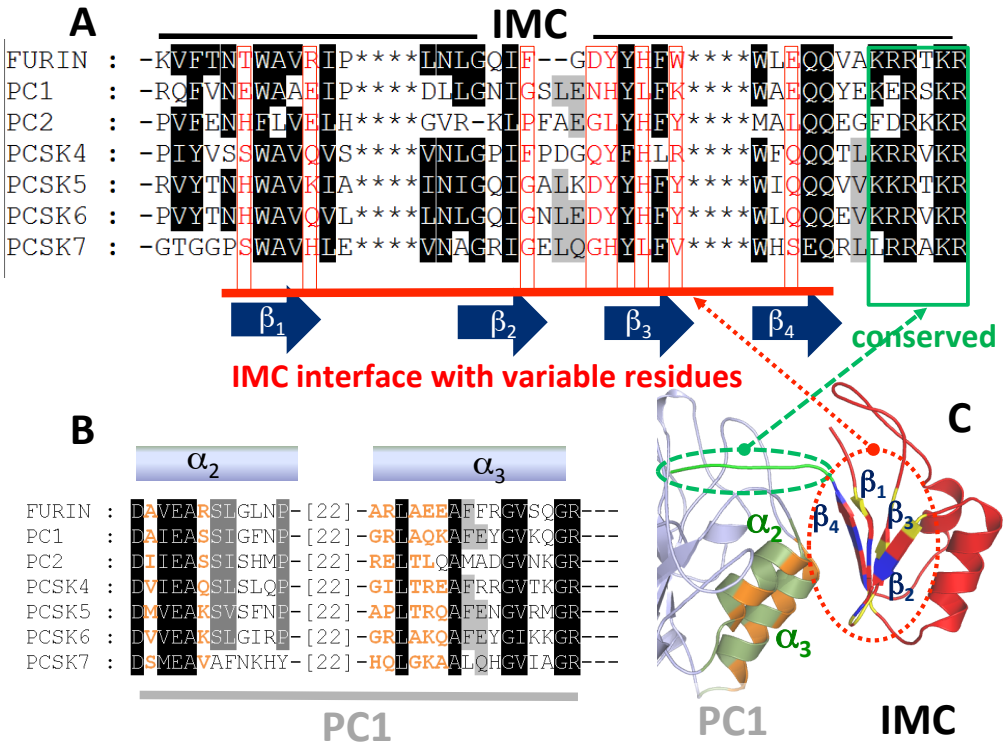


Figure 6.2: Mapping the propeptide-protease interface in PCs.
(A) Highly conserved (black), conserved (grey) and highly variable (red) residues in propeptides (IMCs) across PCs. Positions marked with stars depict residues not shown. **(B)** Conserved residues in an interface comprised of two helices stacked against the beta sheets of the propeptide. Highly conserved residues are black, conserved are grey and variable are orange. The 22-residue stretch between helices has been deleted. **(C)** Ribbon diagram of the binding interface between IMC and protease.

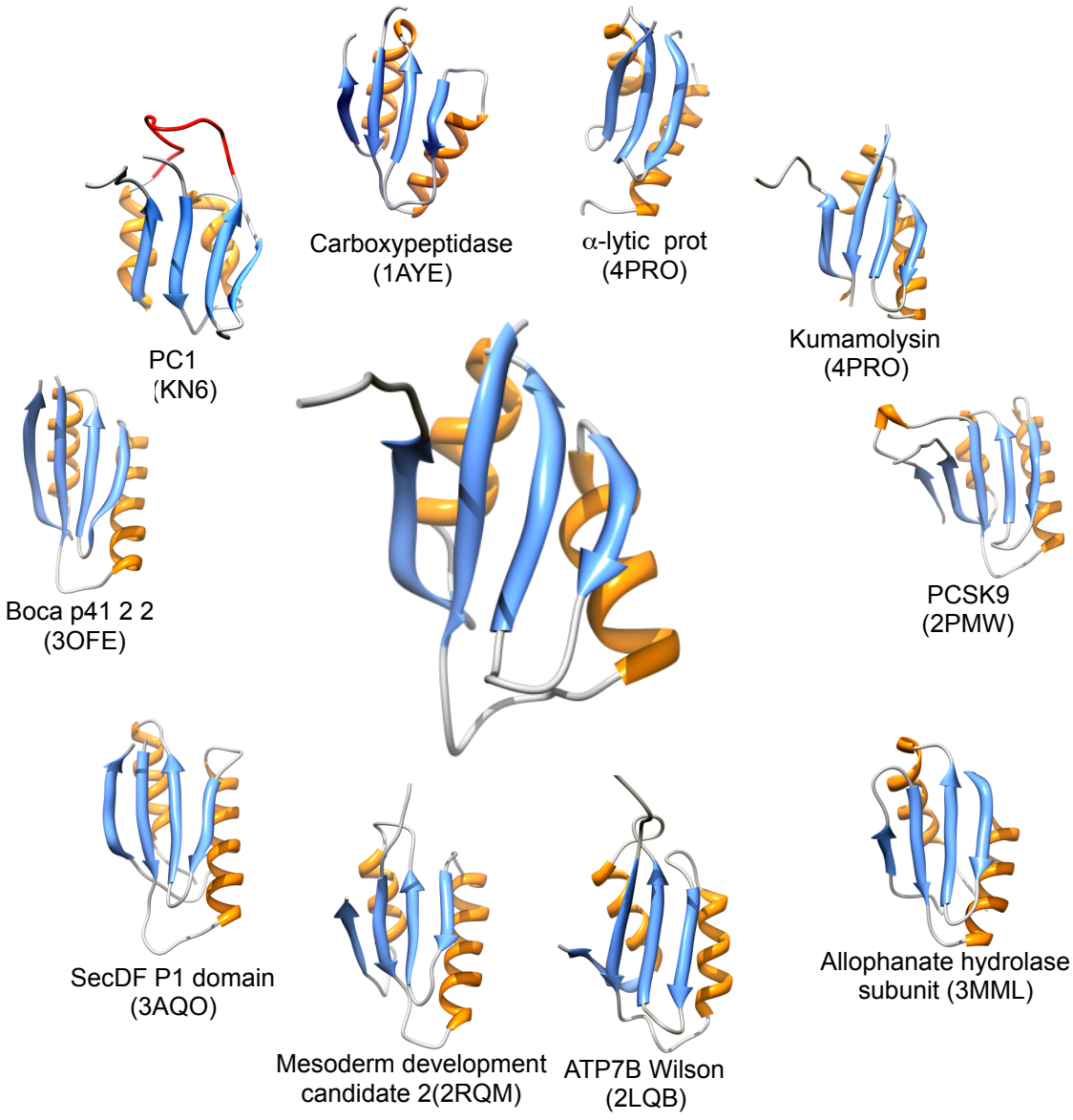


Figure 6.3: Conserved ferredoxin fold across diverse proteins
The structure of the conserved ferredoxin fold is represented in the center. Select examples of proteins containing this conserved fold are given in the periphery, with PDB identifiers.

References

1. Steiner, D.F., et al., *Insulin biosynthesis: evidence for a precursor.* Science, 1967. **157**(3789): p. 697-700.
2. Rawlings, N.D., A.J. Barrett, and A. Bateman, *MEROPS: the peptidase database.* Nucleic Acids Res, 2010. **38**(Database issue): p. D227-33.
3. Siezen, R.J. and J.A. Leunissen, *Subtilases: the superfamily of subtilisin-like serine proteases.* Protein Sci, 1997. **6**(3): p. 501-23.
4. Bryan, P.N., *Protein engineering of subtilisin.* Biochim Biophys Acta, 2000. **1543**(2): p. 203-222.
5. Shinde, U. and M. Inouye, *Intramolecular chaperones: polypeptide extensions that modulate protein folding.* Semin Cell Dev Biol, 2000. **11**(1): p. 35-44.
6. Henrich, S., et al., *The crystal structure of the proprotein processing proteinase furin explains its stringent specificity.* Nat Struct Biol, 2003. **10**(7): p. 520-6.
7. Holyoak, T., et al., *2.4 A resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg boronic acid inhibitor.* Biochemistry, 2003. **42**(22): p. 6709-18.
8. Oliva, A.A., Jr., D.F. Steiner, and S.J. Chan, *Proprotein convertases in amphioxus: predicted structure and expression of proteases SPC2 and SPC3.* Proc Natl Acad Sci U S A, 1995. **92**(8): p. 3591-5.
9. Lipkind, G., Q. Gong, and D.F. Steiner, *Molecular modeling of the substrate specificity of prohormone convertases SPC2 and SPC3.* J Biol Chem, 1995. **270**(22): p. 13277-84.
10. Henrich, S., et al., *Proprotein convertase models based on the crystal structures of furin and kexin: explanation of their specificity.* J Mol Biol, 2005. **345**(2): p. 211-27.

11. Bergeron, F., R. Leduc, and R. Day, *Subtilase-like pro-protein convertases: from molecular specificity to therapeutic applications.* J Mol Endocrinol, 2000. **24**(1): p. 1-22.

12. Seidah, N.G., et al., *The secretory proprotein convertase neural apoptosis-regulated convertase 1 (NARC-1): liver regeneration and neuronal differentiation.* Proc Natl Acad Sci U S A, 2003. **100**(3): p. 928-33.

13. Seidah, N.G., et al., *The activation and physiological functions of the proprotein convertases.* Int J Biochem Cell Biol, 2008. **40**(6-7): p. 1111-25.

14. Seidah, N.G., et al., *Mammalian subtilisin/kexin isozyme SKI-1: A widely expressed proprotein convertase with a unique cleavage specificity and cellular localization.* Proc Natl Acad Sci U S A, 1999. **96**(4): p. 1321-6.

15. Toure, B.B., et al., *Biosynthesis and enzymatic characterization of human SKI-1/S1P and the processing of its inhibitory prosegment.* J Biol Chem, 2000. **275**(4): p. 2349-58.

16. Hampton, E.N., et al., *The self-inhibited structure of full-length PCSK9 at 1.9 A reveals structural homology with resistin within the C-terminal domain.* Proc Natl Acad Sci U S A, 2007. **104**(37): p. 14604-9.

17. Mei, H.C., et al., *Engineering subtilisin YaB: restriction of substrate specificity by the substitution of Gly124 and Gly151 with Ala.* Protein Eng, 1998. **11**(2): p. 109-17.

18. Ikemura, H., H. Takagi, and M. Inouye, *Requirement of pro-sequence for the production of active subtilisin E in Escherichia coli.* J Biol Chem, 1987. **262**(16): p. 7859-64.

19. Ikemura, H. and M. Inouye, *In vitro processing of pro-subtilisin produced in Escherichia coli.* J Biol Chem, 1988. **263**(26): p. 12959-63.

20. Zhu, X.L., et al., *Pro-sequence of subtilisin can guide the refolding of denatured subtilisin in an intermolecular process.* Nature, 1989. **339**(6224): p. 483-4.

21. Eder, J. and A.R. Fersht, *Pro-sequence-assisted protein folding.* Mol Microbiol, 1995. **16**(4): p. 609-14.

22. Chen, Y.J. and M. Inouye, *The intramolecular chaperone-mediated protein folding.* Curr Opin Struct Biol, 2008. **18**(6): p. 765-70.

23. Silen, J.L. and D.A. Agard, *The alpha-lytic protease pro-region does not require a physical linkage to activate the protease domain in vivo.* Nature, 1989. **341**(6241): p. 462-4.

24. Baker, D., J.L. Sohl, and D.A. Agard, *A protein-folding reaction under kinetic control.* Nature, 1992. **356**(6366): p. 263-5.

25. Jaswal, S.S., et al., *Energetic landscape of alpha-lytic protease optimizes longevity through kinetic stability.* Nature, 2002. **415**(6869): p. 343-6.

26. McIver, K.S., E. Kessler, and D.E. Ohman, *Identification of residues in the Pseudomonas aeruginosa elastase propeptide required for chaperone and secretion activities.* Microbiology, 2004. **150**(Pt 12): p. 3969-77.

27. Tang, B., et al., *General function of N-terminal propeptide on assisting protein folding and inhibiting catalytic activity based on observations with a chimeric thermolysin-like protease.* Biochem Biophys Res Commun, 2003. **301**(4): p. 1093-8.

28. Valls, L.A., et al., *Protein sorting in yeast: the localization determinant of yeast vacuolar carboxypeptidase Y resides in the propeptide.* Cell, 1987. **48**(5): p. 887-97.

29. Ramos, C., J.R. Winther, and M.C. Kielland-Brandt, *Requirement of the propeptide for in vivo formation of active yeast carboxypeptidase Y.* J Biol Chem, 1994. **269**(9): p. 7006-12.

30. Winther, J.R. and P. Sorensen, *Propeptide of carboxypeptidase Y provides a chaperone-like function as well as inhibition of the enzymatic activity.* Proc Natl Acad Sci U S A, 1991. **88**(20): p. 9330-4.

31. Winther, J.R., P. Sorensen, and M.C. Kielland-Brandt, *Refolding of a carboxypeptidase Y folding intermediate in vitro by low-affinity binding of the proregion.* J Biol Chem, 1994. **269**(35): p. 22007-13.

32. van den Hazel, H.B., M.C. Kielland-Brandt, and J.R. Winther, *The propeptide is required for in vivo formation of stable active yeast proteinase A and can function even when not covalently linked to the mature region.* J Biol Chem, 1993. **268**(24): p. 18002-7.

33. Smith, S.M. and M.M. Gottesman, *Activity and deletion analysis of recombinant human cathepsin L expressed in Escherichia coli.* J Biol Chem, 1989. **264**(34): p. 20487-95.

34. Ogino, T., et al., *Function of the propeptide region in recombinant expression of active procathepsin L in Escherichia coli.* J Biochem, 1999. **126**(1): p. 78-83.

35. Wiederanders, B., G. Kaulmann, and K. Schilling, *Functions of propeptide parts in cysteine proteases.* Curr Protein Pept Sci, 2003. **4**(5): p. 309-26.

36. Schulz, E.C., et al., *Structure analysis of endosialidase NF at 0.98 A resolution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 176-80.

37. Schulz, E.C., et al., *Structural basis for the recognition and cleavage of polysialic acid by the bacteriophage K1F tailspike protein EndoNF.* J Mol Biol, 2010. **397**(1): p. 341-51.

38. Shinde, U. and M. Inouye, *Intramolecular chaperones and protein folding.* Trends Biochem Sci, 1993. **18**(11): p. 442-6.

39. Ellis, R.J., *The general concept of molecular chaperones.* Philos Trans R Soc Lond B Biol Sci, 1993. **339**(1289): p. 257-61.

40. Ellis, R.J., *Protein misassembly: macromolecular crowding and molecular chaperones.* Adv Exp Med Biol, 2007. **594**: p. 1-13.

41. Ellis, R.J. and S.M. van der Vies, *Molecular chaperones.* Annu Rev Biochem, 1991. **60**: p. 321-47.

42. Shinde, U., et al., *Folding pathway mediated by an intramolecular chaperone.* Proc Natl Acad Sci U S A, 1993. **90**(15): p. 6924-8.

43. Eder, J., M. Rheinnecker, and A.R. Fersht, *Folding of subtilisin BPN': role of the pro-sequence.* J Mol Biol, 1993. **233**(2): p. 293-304.

44. Eder, J., M. Rheinnecker, and A.R. Fersht, *Folding of subtilisin BPN': characterization of a folding intermediate.* Biochemistry, 1993. **32**(1): p. 18-26.

45. Bryan, P., et al., *Energetics of folding subtilisin BPN'.* Biochemistry, 1992. **31**(21): p. 4937-45.

46. Gallagher, T., et al., *The prosegment-subtilisin BPN' complex: crystal structure of a specific 'foldase'.* Structure, 1995. **3**(9): p. 907-14.

47. Jain, S.C., et al., *The crystal structure of an autoprocessed Ser221Cys-subtilisin E-propeptide complex at 2.0 A resolution.* J Mol Biol, 1998. **284**(1): p. 137-44.

48. Shinde, U., X. Fu, and M. Inouye, *A pathway for conformational diversity in proteins mediated by intramolecular chaperones.* J Biol Chem, 1999. **274**(22): p. 15615-21.

49. Yabuta, Y., et al., *Folding pathway mediated by an intramolecular chaperone: propeptide release modulates activation precision of pro-subtilisin.* J Biol Chem, 2001. **276**(48): p. 44427-34.

50. Ohta, Y., et al., *Pro-peptide as an intramolecular chaperone: renaturation of denatured subtilisin E with a synthetic pro-peptide [corrected].* Mol Microbiol, 1991. **5**(6): p. 1507-10.

51. Ohta, Y. and M. Inouye, *Pro-subtilisin E: purification and characterization of its autoprocessing to active subtilisin E in vitro.* Mol Microbiol, 1990. **4**(2): p. 295-304.

52. Baker, D. and D.A. Agard, *Kinetics versus thermodynamics in protein folding.* Biochemistry, 1994. **33**(24): p. 7505-9.

53. Shinde, U. and M. Inouye, *The structural and functional organization of intramolecular chaperones: the N-terminal propeptides which mediate protein folding.* J Biochem, 1994. **115**(4): p. 629-36.

54. Kojima, S., A. Iwahara, and H. Yanai, *Inhibitor-assisted refolding of protease: a protease inhibitor as an intramolecular chaperone.* FEBS Lett, 2005. **579**(20): p. 4430-6.

55. Yabuta, Y., et al., *Folding pathway mediated by an intramolecular chaperone. A functional peptide chaperone designed using sequence databases.* J Biol Chem, 2003. **278**(17): p. 15246-51.

56. Li, Y., et al., *Functional analysis of the propeptide of subtilisin E as an intramolecular chaperone for protein folding. Refolding and inhibitory abilities of propeptide mutants.* J Biol Chem, 1995. **270**(42): p. 25127-32.

57. Kobayashi, T. and M. Inouye, *Functional analysis of the intramolecular chaperone. Mutational hot spots in the subtilisin pro-peptide and a second-site suppressor mutation within the subtilisin molecule.* J Mol Biol, 1992. **226**(4): p. 931-3.

58. Radisky, E.S., et al., *The role of the protein core in the inhibitory power of the classic serine protease inhibitor, chymotrypsin inhibitor 2.* Biochemistry, 2003. **42**(21): p. 6484-92.

59. Bryan, P., et al., *Catalysis of a protein folding reaction: mechanistic implications of the 2.0 A structure of the subtilisin-prodomain complex.* Biochemistry, 1995. **34**(32): p. 10310-8.

60. Baker, D., A.K. Shiau, and D.A. Agard, *The role of pro regions in protein folding.* Curr Opin Cell Biol, 1993. **5**(6): p. 966-70.

61. Kuwajima, K., *The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure.* Proteins, 1989. **6**(2): p. 87-103.

62. Sohl, J.L., S.S. Jaswal, and D.A. Agard, *Unfolded conformations of alpha-lytic protease are more stable than its native state.* Nature, 1998. **395**(6704): p. 817-9.

63. Dobson, C.M., *Principles of protein folding, misfolding and aggregation.* Semin Cell Dev Biol, 2004. **15**(1): p. 3-16.

64. Dobson, C.M., *Protein folding and misfolding.* Nature, 2003. **426**(6968): p. 884-90.

65. Subbian, E., Y. Yabuta, and U. Shinde, *Positive selection dictates the choice between kinetic and thermodynamic protein folding and stability in subtilases.* Biochemistry, 2004. **43**(45): p. 14348-60.

66. Subbian, E., Y. Yabuta, and U.P. Shinde, *Folding pathway mediated by an intramolecular chaperone: intrinsically unstructured propeptide modulates stochastic activation of subtilisin.* J Mol Biol, 2005. **347**(2): p. 367-83.

67. Plaxco, K.W., et al., *Evolutionary conservation in protein folding kinetics.* J Mol Biol, 2000. **298**(2): p. 303-12.

68. Nishimura, C., et al., *Conservation of folding pathways in evolutionarily distant globin sequences.* Nat Struct Biol, 2000. **7**(8): p. 679-86.

69. Shakhnovich, E., V. Abkevich, and O. Ptitsyn, *Conserved residues and the mechanism of protein folding.* Nature, 1996. **379**(6560): p. 96-8.

70. Kucerova, H., et al., *Differences in the regulation of the intracellular Ca2+-dependent serine proteinase activity between Bacillus subtilis and B. megaterium.* Curr Microbiol, 2001. **42**(3): p. 178-83.

71. Vachova, L., *Activation of the intracellular Ca(2+)-dependent serine protease ISP1 of bacillus megaterium by purification or by high Ca2+ concentrations.* Biochem Mol Biol Int, 1996. **40**(5): p. 947-54.

72. Vachova, L., et al., *Heat and osmotic stress enhance the development of cytoplasmic serine proteinase activity in sporulating Bacillus megaterium.* Biochem Mol Biol Int, 1994. **32**(6): p. 1049-57.

73. Kucerova, H. and J. Chaloupka, *Intracellular serine proteinase behaves as a heat-stress protein in nongrowing but as a cold-stress protein in growing populations of Bacillus megaterium.* Curr Microbiol, 1995. **31**(1): p. 39-43.

74. Shinde, U. and M. Inouye, *Folding pathway mediated by an intramolecular chaperone: characterization of the structural changes in pro-subtilisin E coincident with autoprocessing.* J Mol Biol, 1995. **252**(1): p. 25-30.

75. Weisz, O.A., *Organelle acidification and disease.* Traffic, 2003. **4**(2): p. 57-64.

76. Fu, X., M. Inouye, and U. Shinde, *Folding pathway mediated by an intramolecular chaperone. The inhibitory and chaperone functions of the subtilisin propeptide are not obligatorily linked.* J Biol Chem, 2000. **275**(22): p. 16871-8.

77. Inouye, M., X. Fu, and U. Shinde, *Substrate-induced activation of a trapped IMC-mediated protein folding intermediate.* Nat Struct Biol, 2001. **8**(4): p. 321-5.

78. McNutt, M.C., T.A. Lagace, and J.D. Horton, *Catalytic activity is not required for secreted PCSK9 to reduce low density lipoprotein receptors in HepG2 cells.* J Biol Chem, 2007. **282**(29): p. 20799-803.

79. Pavlaki, M., et al., *Furin Functions as a Nonproteolytic Chaperone for Matrix Metalloproteinase-28: MMP-28 Propeptide Sequence Requirement.* Biochem Res Int, 2011. **2011**: p. 630319.

80. Fuller, R.S., A.J. Brake, and J. Thorner, *Intracellular targeting and structural conservation of a prohormone-processing endoprotease.* Science, 1989. **246**(4929): p. 482-6.

81. Nour, N., et al., *Structure-function analysis of the prosegment of the proprotein convertase PC5A.* J Biol Chem, 2003. **278**(5): p. 2886-95.

82. Ueda, K., et al., *Mutational analysis of predicted interactions between the catalytic and P domains of prohormone convertase 3 (PC3/PC1).* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5622-7.

83. Zhou, A., et al., *Regulatory roles of the P domain of the subtilisin-like prohormone convertases.* J Biol Chem, 1998. **273**(18): p. 11107-14.

84. Comellas-Bigler, M., et al., *Crystal structure of the E. coli dipeptidyl carboxypeptidase Dcp: further indication of a ligand-dependent hinge movement mechanism.* J Mol Biol, 2005. **349**(1): p. 99-112.

85. Dikeakos, J.D., et al., *A hydrophobic patch in a charged alpha-helix is sufficient to target proteins to dense core secretory granules.* J Biol Chem, 2007. **282**(2): p. 1136-43.

86. Dikeakos, J.D., et al., *Functional and structural characterization of a dense core secretory granule sorting domain from the PC1/3 protease.* Proc Natl Acad Sci U S A, 2009. **106**(18): p. 7408-13.

87. Rehemtulla, A., A.J. Dorner, and R.J. Kaufman, *Regulation of PACE propeptide-processing activity: requirement for a post-endoplasmic reticulum compartment and autoproteolytic activation.* Proc Natl Acad Sci U S A, 1992. **89**(17): p. 8235-9.

88. Thomas, G., *Furin at the cutting edge: from protein traffic to embryogenesis and disease.* Nat Rev Mol Cell Biol, 2002. **3**(10): p. 753-66.

89. Anderson, E.D., et al., *Activation of the furin endoprotease is a multiple-step process: requirements for acidification and internal propeptide cleavage.* EMBO J, 1997. **16**(7): p. 1508-18.

90. Tangrea, M.A., et al., *Stability and global fold of the mouse prohormone convertase 1 pro-domain.* Biochemistry, 2001. **40**(18): p. 5488-95.

91. Tangrea, M.A., et al., *Solution structure of the pro-hormone convertase 1 pro-domain from Mus musculus.* J Mol Biol, 2002. **320**(4): p. 801-12.

92. Dillon, S.L., et al., *Propeptides are sufficient to regulate organelle-specific pH-dependent activation of furin and proprotein convertase 1/3.* J Mol Biol, 2012. **423**(1): p. 47-62.

93. Leduc, R., et al., *Activation of human furin precursor processing endoprotease occurs by an intramolecular autoproteolytic cleavage.* J Biol Chem, 1992. **267**(20): p. 14304-8.

94. Anderson, E.D., et al., *The ordered and compartment-specfific autoproteolytic removal of the furin intramolecular chaperone is required for enzyme activation.* J Biol Chem, 2002. **277**(15): p. 12879-90.

95. Elferich, J., et al., *Propeptides of eukaryotic proteases encode histidines to exploit organelle pH for regulation.* FASEB J, 2013. **27**(8): p. 2939-45.

96. Demaurex, N., *pH Homeostasis of cellular organelles.* News Physiol Sci, 2002. **17**: p. 1-5.

97. Casey, J.R., S. Grinstein, and J. Orlowski, *Sensors and regulators of intracellular pH.* Nat Rev Mol Cell Biol, 2010. **11**(1): p. 50-61.

98. Rotzschke, O., et al., *A pH-sensitive histidine residue as control element for ligand release from HLA-DR molecules.* Proc Natl Acad Sci U S A, 2002. **99**(26): p. 16946-50.

99. Perutz, M.F., *Electrostatic effects in proteins.* Science, 1978. **201**(4362): p. 1187-91.

100. Perier, A., et al., *Concerted protonation of key histidines triggers membrane interaction of the diphtheria toxin T domain.* J Biol Chem, 2007. **282**(33): p. 24239-45.

101. Schmidt, W.K. and H.P. Moore, *Ionic milieu controls the compartment-specific activation of pro-opiomelanocortin processing in AtT-20 cells.* Mol Biol Cell, 1995. **6**(10): p. 1271-85.

102. Clapham, D.E., *Calcium signaling.* Cell, 2007. **131**(6): p. 1047-58.

103. Feliciangeli, S.F., et al., *Identification of a pH sensor in the furin propeptide that regulates enzyme activation.* J Biol Chem, 2006. **281**(23): p. 16108-16.

104. Williamson, D.M., et al., *The mechanism by which a propeptide-encoded pH sensor regulates spatiotemporal activation of furin.* J Biol Chem, 2013. **288**(26): p. 19154-65.

105. Williamson, D.M., J. Elferich, and U. Shinde, *Mechanism of Fine-tuning pH Sensors in Proprotein Convertases: Identification of a pH-sensing Histidine Pair in the Propeptide of Proprotein Convertase 1/3.* Journal of Biological Chemistry, 2015.

106. Elferich, J., et al., *Determination of Histidine pKa Values in the Propeptides of Furin and PC1/3 using Histidine Hydrogen-Deuterium Exchange Mass Spectrometry.* Anal Chem, 2015.

107. Matsuyama, S., et al., *Changes in intramitochondrial and cytosolic pH: early events that modulate caspase activation during apoptosis.* Nat Cell Biol, 2000. **2**(6): p. 318-25.

108. Schreiber, R., *Ca2+ signaling, intracellular pH and cell volume in cell proliferation.* J Membr Biol, 2005. **205**(3): p. 129-37.

109. Webb, B.A., et al., *Dysregulated pH: a perfect storm for cancer progression.* Nat Rev Cancer, 2011. **11**(9): p. 671-7.

110. Naghavi, M., et al., *pH Heterogeneity of human and rabbit atherosclerotic plaques; a new insight into detection of vulnerable plaque.* Atherosclerosis, 2002. **164**(1): p. 27-35.

111. Lucien, F., et al., *Hypoxia-induced invadopodia formation involves activation of NHE-1 by the p90 ribosomal S6 kinase (p90RSK).* PLoS One, 2011. **6**(12): p. e28851.

112. Shinde, U.P., J.J. Liu, and M. Inouye, *Protein memory through altered folding mediated by intramolecular chaperones.* Nature, 1997. **389**(6650): p. 520-2.

113. Pickett, L.A., et al., *Functional consequences of a novel variant of PCSK1.* PLoS One, 2013. **8**(1): p. e55065.

114. Shinde, U. and M. Inouye, *Folding mediated by an intramolecular chaperone: autoprocessing pathway of the precursor resolved via a substrate assisted catalysis mechanism.* J Mol Biol, 1995. **247**(3): p. 390-5.

115. Seidah, N.G., *The proprotein convertases, 20 years later.* Methods Mol Biol, 2011. **768**: p. 23-57.

116. Seidah, N.G., *What lies ahead for the proprotein convertases?* Ann N Y Acad Sci, 2011. **1220**: p. 149-61.

117. Fuller, R.S., A. Brake, and J. Thorner, *Yeast prohormone processing enzyme (KEX2 gene product) is a Ca2+-dependent serine protease.* Proc Natl Acad Sci U S A, 1989. **86**(5): p. 1434-8.

118. Fuller, R.S., R.E. Sterne, and J. Thorner, *Enzymes required for yeast prohormone processing.* Annu Rev Physiol, 1988. **50**: p. 345-62.

119. Thomas, G., et al., *Yeast KEX2 endopeptidase correctly cleaves a neuroendocrine prohormone in mammalian cells.* Science, 1988. **241**(4862): p. 226-30.

120. Holyoak, T., et al., *Structural basis for differences in substrate selectivity in Kex2 and furin protein convertases.* Biochemistry, 2004. **43**(9): p. 2412-21.

121. Bard, F. and V. Malhotra, *The formation of TGN-to-plasma-membrane transport carriers.* Annu Rev Cell Dev Biol, 2006. **22**: p. 439-55.

122. Shinde, U. and G. Thomas, *Insights from bacterial subtilases into the mechanisms of intramolecular chaperone-mediated activation of furin.* Methods Mol Biol, 2011. **768**: p. 59-106.

123. Bassi, D.E., et al., *Increased furin activity enhances the malignant phenotype of human head and neck cancer cells.* Am J Pathol, 2003. **162**(2): p. 439-47.

124. Bassi, D.E., H. Mahloogi, and A.J. Klein-Szanto, *The proprotein convertases furin and PACE4 play a significant role in tumor progression.* Mol Carcinog, 2000. **28**(2): p. 63-9.

125. Choquet, H., P. Stijnen, and J.W. Creemers, *Genetic and functional characterization of PCSK1.* Methods Mol Biol, 2011. **768**: p. 247-53.

126. Benzinou, M., et al., *Common nonsynonymous variants in PCSK1 confer risk of obesity.* Nat Genet, 2008. **40**(8): p. 943-5.

127. Steiner, D.F., et al., *The role of prohormone convertases in insulin biosynthesis: evidence for inherited defects in their action in man and experimental animals.* Diabetes Metab, 1996. **22**(2): p. 94-104.

128. Ehret, G.B., et al., *Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk.* Nature, 2011. **478**(7367): p. 103-9.

129. Bassi, D.E., et al., *Furin inhibition results in absent or decreased invasiveness and tumorigenicity of human cancer cells.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10326-31.

130. Inouye, M., *Intramolecular chaperone: the role of the pro-peptide in protein folding.* Enzyme, 1991. **45**(5-6): p. 314-21.

131. Siezen, R.J., *Subtilases: subtilisin-like serine proteases.* Adv Exp Med Biol, 1996. **379**: p. 75-93.

132. Siezen, R.J., J.W. Creemers, and W.J. Van de Ven, *Homology modelling of the catalytic domain of human furin. A model for the eukaryotic subtilisin-like proprotein convertases.* Eur J Biochem, 1994. **222**(2): p. 255-66.

133. Cunningham, E.L., et al., *Kinetic stability as a mechanism for protease longevity.* Proc Natl Acad Sci U S A, 1999. **96**(20): p. 11008-14.

134. Truhlar, S.M., E.L. Cunningham, and D.A. Agard, *The folding landscape of Streptomyces griseus protease B reveals the energetic costs and benefits*

associated with evolving kinetic stability. Protein Sci, 2004. **13**(2): p. 381-90.

135. Nakayama, K., *Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins.* Biochem J, 1997. **327 (Pt 3)**: p. 625-35.

136. Embley, T.M. and W. Martin, *Eukaryotic evolution, changes and challenges.* Nature, 2006. **440**(7084): p. 623-30.

137. Soskine, M. and D.S. Tawfik, *Mutational effects and the evolution of new protein functions.* Nat Rev Genet, 2010. **11**(8): p. 572-82.

138. Marie-Claire, C., et al., *Folding pathway mediated by an intramolecular chaperone: the structural and functional characterization of the aqualysin I propeptide.* J Mol Biol, 2001. **305**(1): p. 151-65.

139. Shinde, U. and G. Thomas, *Insights from bacterial subtilases into the mechanisms of intramolecular chaperone mediated activation of furin.* Methods in Molecular Biology, 2011. **(in press)**.

140. Bhattacharjya, S., et al., *pH-induced conformational transitions of a molten-globule-like state of the inhibitory prodomain of furin: implications for zymogen activation.* Protein Sci, 2001. **10**(5): p. 934-42.

141. Fink, A.L., et al., *Classification of acid denaturation of proteins: intermediates and unfolded states.* Biochemistry, 1994. **33**(41): p. 12504-11.

142. Goto, Y., L.J. Calciano, and A.L. Fink, *Acid-induced folding of proteins.* Proc Natl Acad Sci U S A, 1990. **87**(2): p. 573-7.

143. Uversky, V.N. and Y. Goto, *Acid denaturation and anion-induced folding of globular proteins: multitude of equilibium partially folded intermediates.* Curr Protein Pept Sci, 2009. **10**(5): p. 447-55.

144. Wright, P.E. and H.J. Dyson, *Linking folding and binding.* Curr Opin Struct Biol, 2009. **19**(1): p. 31-8.

145. Gething, M.J., K. McCammon, and J. Sambrook, *Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport.* Cell, 1986. **46**(6): p. 939-50.

146. Molloy, S.S., et al., *Intracellular trafficking and activation of the furin proprotein convertase: localization to the TGN and recycling from the cell surface.* Embo J, 1994. **13**(1): p. 18-33.

147. Munro, S. and H.R. Pelham, *A C-terminal signal prevents secretion of luminal ER proteins.* Cell, 1987. **48**(5): p. 899-907.

148. Karplus, M. and J. Kuriyan, *Molecular dynamics and protein function.* Proc Natl Acad Sci U S A, 2005. **102**(19): p. 6679-85.

149. Daggett, V. and M. Levitt, *A model of the molten globule state from molecular dynamics simulations.* Proc Natl Acad Sci U S A, 1992. **89**(11): p. 5142-6.

150. Daggett, V. and M. Levitt, *Protein unfolding pathways explored through molecular dynamics simulations.* J Mol Biol, 1993. **232**(2): p. 600-19.

151. Brooks, C.L., 3rd, *Characterization of "native" apomyoglobin by molecular dynamics simulation.* J Mol Biol, 1992. **227**(2): p. 375-80.

152. Salimi, N.L., B. Ho, and D.A. Agard, *Unfolding simulations reveal the mechanism of extreme unfolding cooperativity in the kinetically stable alpha-lytic protease.* PLoS Comput Biol, 2010. **6**(2): p. e1000689.

153. Huang, W., A.P. Eichenberger, and W.F. van Gunsteren, *Molecular dynamics simulation of thionated hen egg white lysozyme.* Protein Sci, 2012.

154. Povolotskaya, I.S. and F.A. Kondrashov, *Sequence space and the ongoing expansion of the protein universe.* Nature, 2010. **465**(7300): p. 922-6.

155. Benjannet, S., et al., *PC1 and PC2 are proprotein convertases capable of cleaving proopiomelanocortin at distinct pairs of basic residues.* Proc Natl Acad Sci U S A, 1991. **88**(9): p. 3564-8.

156. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nat Protoc, 2006. **1**(6): p. 2876-90.

157. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD.* J Comput Chem, 2005. **26**(16): p. 1781-802.

158. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program.* J Comput Chem, 2009. **30**(10): p. 1545-614.

159. Roux, B., B. Prod'hom, and M. Karplus, *Ion transport in the gramicidin channel: molecular dynamics study of single and double occupancy.* Biophys J, 1995. **68**(3): p. 876-92.

160. Chretien, M., et al., *Proprotein convertases as therapeutic targets.* Expert Opin Ther Targets, 2008. **12**(10): p. 1289-300.

161. Artenstein, A.W. and S.M. Opal, *Proprotein convertases in health and disease.* N Engl J Med, 2011. **365**(26): p. 2507-18.

162. Scamuffa, N., et al., *Proprotein convertases: lessons from knockouts.* FASEB J, 2006. **20**(12): p. 1954-63.

163. Kim, W., et al., *Loss of endothelial furin leads to cardiac malformation and early postnatal death.* Mol Cell Biol, 2012. **32**(17): p. 3382-91.

164. Creemers, J.W., et al., *Endoproteolytic cleavage of its propeptide is a prerequisite for efficient transport of furin out of the endoplasmic reticulum.* J Biol Chem, 1995. **270**(6): p. 2695-702.

165. Vey, M., et al., *Maturation of the trans-Golgi network protease furin: compartmentalization of propeptide removal, substrate cleavage, and COOH-terminal truncation.* J Cell Biol, 1994. **127**(6 Pt 2): p. 1829-42.

166. Elferich, J., et al., *Propeptides of eukaryotic proteases encode histidines to exploit organelle pH for regulation.* FASEB J, 2013.

167. Dhar, P., et al., *Cellware--a multi-algorithmic software for computational systems biology.* Bioinformatics, 2004. **20**(8): p. 1319-21.

168. Greenfield, N.J., *Analysis of the kinetics of folding of proteins and peptides using circular dichroism.* Nat Protoc, 2006. **1**(6): p. 2891-9.

169. Nair, S.K., et al., *Conformational dynamics of estrogen receptors alpha and beta as revealed by intrinsic tryptophan fluorescence and circular dichroism.* J Mol Endocrinol, 2005. **35**(2): p. 211-23.

170. Lakowicz, J.R., et al., *Rotational freedom of tryptophan residues in proteins and peptides.* Biochemistry, 1983. **22**(8): p. 1741-52.

171. Greenfield, N.J., *Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism.* Nat Protoc, 2006. **1**(6): p. 2733-41.

172. Sawicki, C.A. and Q.H. Gibson, *Quaternary conformational changes in human hemoglobin studied by laser photolysis of carboxyhemoglobin.* J Biol Chem, 1976. **251**(6): p. 1533-42.

173. Bhattacharjya, S., et al., *Conformational analyses of a partially-folded bioactive prodomain of human furin.* Biopolymers, 2007. **86**(4): p. 329-44.

174. Basak, A. and C. Lazure, *Synthetic peptides derived from the prosegments of proprotein convertase 1/3 and furin are potent inhibitors of both enzymes.* Biochem J, 2003. **373**(Pt 1): p. 231-9.

175. Jean, F., et al., *Enzymic characterization of murine and human prohormone convertase-1 (mPC1 and hPC1) expressed in mammalian GH4C1 cells.* Biochem J, 1993. **292 (Pt 3)**: p. 891-900.

176. Dufour, E.K., et al., *Stability of mutant serpin/furin complexes: dependence on pH and regulation at the deacylation step.* Protein Sci, 2005. **14**(2): p. 303-15.

177. Huang, W., A.P. Eichenberger, and W.F. van Gunsteren, *Molecular dynamics simulation of thionated hen egg white lysozyme.* Protein Sci, 2012. **21**(8): p. 1153-61.

178. Basak, A., et al., *Design and synthesis of novel inhibitors of prohormone convertases.* Int J Pept Protein Res, 1994. **44**(3): p. 253-61.

179. Day, R., et al., *Prodynorphin processing by proprotein convertase 2. Cleavage at single basic residues and enhanced processing in the presence of carboxypeptidase activity.* J Biol Chem, 1998. **273**(2): p. 829-36.

180. Tanco, S., et al., *Characterization of the substrate specificity of human carboxypeptidase A4 and implications for a role in extracellular peptide processing.* J Biol Chem, 2010. **285**(24): p. 18385-96.

181. Levin, L.R. and J. Buck, *Physiological roles of acid-base sensors.* Annu Rev Physiol, 2015. **77**: p. 347-62.

182. Srivastava, J., D.L. Barber, and M.P. Jacobson, *Intracellular pH sensors: design principles and functional significance.* Physiology (Bethesda), 2007. **22**: p. 30-9.

183. Seidah, N.G., et al., *The multifaceted proprotein convertases: their unique, redundant, complementary, and opposite functions.* J Biol Chem, 2013. **288**(30): p. 21473-81.

184. Rashid, S., et al., *Proprotein convertase subtilisin kexin type 9 promotes intestinal overproduction of triglyceride-rich apolipoprotein B lipoproteins through both low-density lipoprotein receptor-dependent and -independent mechanisms.* Circulation, 2014. **130**(5): p. 431-41.

185. Fu, J., et al., *Transgenic overexpression of the proprotein convertase furin enhances skin tumor growth.* Neoplasia, 2012. **14**(4): p. 271-82.

186. Arsenault, D., F. Lucien, and C.M. Dubois, *Hypoxia enhances cancer cell invasion through relocalization of the proprotein convertase furin from the trans-Golgi network to the cell surface.* J Cell Physiol, 2012. **227**(2): p. 789-800.

187. Bassi, D.E., et al., *Proprotein convertase inhibition results in decreased skin cell proliferation, tumorigenesis, and metastasis.* Neoplasia, 2010. **12**(7): p. 516-26.

188. Scamuffa, N., et al., *Prodomain of the proprotein convertase subtilisin/kexin Furin (ppFurin) protects from tumor progression and metastasis.* Carcinogenesis, 2014. **35**(3): p. 528-36.

189. Lopez de Cicco, R., et al., *Human carcinoma cell growth and invasiveness is impaired by the propeptide of the ubiquitous proprotein convertase furin.* Cancer Res, 2005. **65**(10): p. 4162-71.

190. Pasquato, A., et al., *Viral envelope glycoprotein processing by proprotein convertases.* Antiviral Res, 2013. **99**(1): p. 49-60.

191. Seidah, N.G., et al., *PCSK9: a key modulator of cardiovascular health.* Circ Res, 2014. **114**(6): p. 1022-36.

192. Susan-Resiga, D., et al., *Furin is the major processing enzyme of the cardiac-specific growth factor bone morphogenetic protein 10.* J Biol Chem, 2011. **286**(26): p. 22785-94.

193. Roebroek, A.J., et al., *Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase Furin.* Development, 1998. **125**(24): p. 4863-76.

194. Zhu, X., et al., *Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10293-8.

195. Stijnen, P., et al., *The association of common variants in PCSK1 with obesity: a HuGE review and meta-analysis.* Am J Epidemiol, 2014. **180**(11): p. 1051-65.

196. Blanco, E.H., et al., *Biochemical and cell biological properties of the human prohormone convertase 1/3 Ser357Gly mutation: a PC1/3 hypermorph.* Endocrinology, 2014. **155**(9): p. 3434-47.

197. Creemers, J.W., et al., *Heterozygous mutations causing partial prohormone convertase 1 deficiency contribute to human obesity.* Diabetes, 2012. **61**(2): p. 383-90.

198. Prabhu, Y., et al., *Defective transport of the obesity mutant PC1/3 N222D contributes to loss of function.* Endocrinology, 2014. **155**(7): p. 2391-401.

199. Farooqi, I.S., et al., *Hyperphagia and early-onset obesity due to a novel homozygous missense mutation in prohormone convertase 1/3.* J Clin Endocrinol Metab, 2007. **92**(9): p. 3369-73.

200. Benjannet, S., et al., *Comparative biosynthesis, covalent post-translational modifications and efficiency of prosegment cleavage of the prohormone convertases PC1 and PC2: glycosylation, sulphation and identification of the intracellular site of prosegment cleavage of PC1 and PC2.* Biochem J, 1993. **294 (Pt 3)**: p. 735-43.

201. Jackson, R.S., et al., *Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency.* J Clin Invest, 2003. **112**(10): p. 1550-60.

202. Jackson, R.S., et al., *Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene.* Nat Genet, 1997. **16**(3): p. 303-6.

203. Martin, M.G., et al., *Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort.* Gastroenterology, 2013. **145**(1): p. 138-48.

204. Lopez-Otin, C. and J.S. Bond, *Proteases: multifunctional enzymes in life and disease.* J Biol Chem, 2008. **283**(45): p. 30433-7.

205. Elferich, J., Dillon, S and Shinde, U. *Evolutionary insights into mechanisms of organelle specific activation of proprotein convertases; Proceeding of the 6th International Conference on Bioinformatics and Biomedical Engineering, May 17-20 2012 in Shanghai, China*. 2012.

206. Paroutis, P., N. Touret, and S. Grinstein, *The pH of the secretory pathway: measurement, determinants, and regulation.* Physiology (Bethesda), 2004. **19**: p. 207-15.

207. Schonichen, A., et al., *Considering protonation as a posttranslational modification regulating protein structure and function.* Annu Rev Biophys, 2013. **42**: p. 289-314.

208. Zachos, C., et al., *A critical histidine residue within LIMP-2 mediates pH sensitive binding to its ligand beta-glucocerebrosidase.* Traffic, 2012. **13**(8): p. 1113-23.

209. Baird, F.E., et al., *Evidence for allosteric regulation of pH-sensitive System A (SNAT2) and System N (SNAT5) amino acid transporter activity involving a conserved histidine residue.* Biochem J, 2006. **397**(2): p. 369-75.

210. Subbian, E., Williamson, D.M. and Shinde, U., *Protein Folding Mediated by an Intramolecular Chaperone: The Energy Landscape for Unimolecular Pro-Subtilisin E Maturation.* Advances in Bioscience and Biotechnology, 2015. **6**: p. 73-88.

211. Tanford, C. and P.K. De, *The unfolding of beta-lactoglobulin at pH 3 by urea, formamide, and other organic substances.* J Biol Chem, 1961. **236**: p. 1711-5.

212. Reimer, U., et al., *Side-chain effects on peptidyl-prolyl cis/trans isomerisation.* J Mol Biol, 1998. **279**(2): p. 449-60.

213. Texter, F.L., et al., *Intramolecular catalysis of a proline isomerization reaction in the folding of dihydrofolate reductase.* Biochemistry, 1992. **31**(25): p. 5687-91.

214. Reimer, U., et al., *Intramolecular assistance of cis/trans isomerization of the histidine-proline moiety.* Biochemistry, 1997. **36**(45): p. 13802-8.

215. Forbes, S.A., et al., *COSMIC: exploring the world's knowledge of somatic mutations in human cancer.* Nucleic Acids Res, 2015. **43**(Database issue): p. D805-11.

216. Huang, Y.H., et al., *Furin overexpression suppresses tumor growth and predicts a better postoperative disease-free survival in hepatocellular carcinoma.* PLoS One, 2012. **7**(7): p. e40738.

217. Philippe, J., et al., *A nonsense loss-of-function mutation in PCSK1 contributes to dominantly inherited human obesity.* Int J Obes (Lond), 2015. **39**(2): p. 295-302.

218. Rabah, N., et al., *Single amino acid substitution in the PC1/3 propeptide can induce significant modifications of its inhibitory profile toward its cognate enzyme.* J Biol Chem, 2006. **281**(11): p. 7556-67.

219. Turpeinen, H., Z. Ortutay, and M. Pesu, *Genetics of the first seven proprotein convertase enzymes in health and disease.* Curr Genomics, 2013. **14**(7): p. 453-67.

220. Smeekens, S.P., et al., *Proinsulin processing by the subtilisin-related proprotein convertases furin, PC2, and PC3.* Proc Natl Acad Sci U S A, 1992. **89**(18): p. 8822-6.

221. Rouille, Y., et al., *Role of the prohormone convertase PC3 in the processing of proglucagon to glucagon-like peptide 1.* J Biol Chem, 1997. **272**(52): p. 32810-6.

222. Zhu, X., et al., *On the processing of proghrelin to ghrelin.* J Biol Chem, 2006. **281**(50): p. 38867-70.

223. Thomas, L., et al., *Kex2-like endoproteases PC2 and PC3 accurately cleave a model prohormone in mammalian cells: evidence for a common core of neuroendocrine processing enzymes.* Proc Natl Acad Sci U S A, 1991. **88**(12): p. 5297-301.

224. Yourshaw, M., et al., *Exome sequencing finds a novel PCSK1 mutation in a child with generalized malabsorptive diarrhea and diabetes insipidus.* J Pediatr Gastroenterol Nutr, 2013. **57**(6): p. 759-67.

225. Pyun, J.A., et al., *Epistasis between polymorphisms in PCSK1 and DBH is associated with premature ovarian failure.* Menopause, 2014. **21**(11): p. 1249-53.

226. Wei, X., et al., *Genetic variants in PCSK1 gene are associated with the risk of coronary artery disease in type 2 diabetes in a Chinese Han population: a case control study.* PLoS One, 2014. **9**(1): p. e87168.

227. Zhu, X., et al., *Severe block in processing of proinsulin to insulin accompanied by elevation of des-64,65 proinsulin intermediates in islets of mice lacking prohormone convertase 1/3.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10299-304.

228. Lloyd, D.J., S. Bohan, and N. Gekakis, *Obesity, hyperphagia and increased metabolic efficiency in Pc1 mutant mice.* Hum Mol Genet, 2006. **15**(11): p. 1884-93.

229. Sherry, S.T., et al., *dbSNP: the NCBI database of genetic variation.* Nucleic Acids Res, 2001. **29**(1): p. 308-11.

230. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.

231. Ng, P.C. and S. Henikoff, *Predicting deleterious amino acid substitutions.* Genome Res, 2001. **11**(5): p. 863-74.

232. Andreotti, A.H., *Native state proline isomerization: an intrinsic molecular switch.* Biochemistry, 2003. **42**(32): p. 9515-24.

233. Tweedy, N.B., et al., *Structure and energetics of a non-proline cis-peptidyl linkage in a proline-202-->alanine carbonic anhydrase II variant.* Biochemistry, 1993. **32**(41): p. 10944-9.

234. Odefey, C., L.M. Mayr, and F.X. Schmid, *Non-prolyl cis-trans peptide bond isomerization as a rate-determining step in protein unfolding and refolding.* J Mol Biol, 1995. **245**(1): p. 69-78.

235. Vanhove, M., et al., *The rate-limiting step in the folding of the cis-Pro167Thr mutant of TEM-1 beta-lactamase is the trans to cis*

isomerization of a non-proline peptide bond. Proteins, 1996. **25**(1): p. 104-11.

236. Puente, X.S., et al., *Human and mouse proteases: a comparative genomic approach.* Nat Rev Genet, 2003. **4**(7): p. 544-58.

237. Di Cera, E., *Serine proteases.* IUBMB Life, 2009. **61**(5): p. 510-5.

238. Chretien, M. and C.H. Li, *Isolation, purification, and characterization of gamma-lipotropic hormone from sheep pituitary glands.* Can J Biochem, 1967. **45**(7): p. 1163-74.

239. Chretien, M., et al., *From beta-lipotropin to beta-endorphin and 'pro-opio-melanocortin'.* Can J Biochem, 1979. **57**(9): p. 1111-21.

240. van de Ven, W.J., et al., *Furin is a subtilisin-like proprotein processing enzyme in higher eukaryotes.* Mol Biol Rep, 1990. **14**(4): p. 265-75.

241. Bassi, D.E., et al., *Proprotein convertases: "master switches" in the regulation of tumor growth and progression.* Mol Carcinog, 2005. **44**(3): p. 151-61.

242. Khatib, A.M., et al., *Inhibition of proprotein convertases is associated with loss of growth and tumorigenicity of HT-29 human colon carcinoma cells: importance of insulin-like growth factor-1 (IGF-1) receptor processing in IGF-1-mediated functions.* J Biol Chem, 2001. **276**(33): p. 30686-93.

243. Mbikay, M., et al., *Comparative analysis of expression of the proprotein convertases furin, PACE4, PC1 and PC2 in human lung tumours.* Br J Cancer, 1997. **75**(10): p. 1509-14.

244. Khatib, A.M., et al., *Proprotein convertases in tumor progression and malignancy: novel targets in cancer therapy.* Am J Pathol, 2002. **160**(6): p. 1921-35.

245. Abifadel, M., et al., *Mutations and polymorphisms in the proprotein convertase subtilisin kexin 9 (PCSK9) gene in cholesterol metabolism and disease.* Hum Mutat, 2009. **30**(4): p. 520-9.

246. Zaid, A., et al., *Proprotein convertase subtilisin/kexin type 9 (PCSK9): hepatocyte-specific low-density lipoprotein receptor degradation and critical role in mouse liver regeneration.* Hepatology, 2008. **48**(2): p. 646-54.

247. Abifadel, M., et al., *Mutations in PCSK9 cause autosomal dominant hypercholesterolemia.* Nat Genet, 2003. **34**(2): p. 154-6.

248. Rashid, S., et al., *Decreased plasma cholesterol and hypersensitivity to statins in mice lacking Pcsk9.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5374-9.

249. Roubtsova, A., et al., *Circulating proprotein convertase subtilisin/kexin 9 (PCSK9) regulates VLDLR protein and triglyceride accumulation in visceral adipose tissue.* Arterioscler Thromb Vasc Biol, 2011. **31**(4): p. 785-91.

250. Li, N., et al., *Associations between genetic variations in the FURIN gene and hypertension.* BMC Med Genet, 2010. **11**: p. 124.

251. International Consortium for Blood Pressure Genome-Wide Association, S., et al., *Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk.* Nature, 2011. **478**(7367): p. 103-9.

252. Kitamura, K. and K. Tomita, *Proteolytic activation of the epithelial sodium channel and therapeutic application of a serine protease inhibitor for the treatment of salt-sensitive hypertension.* Clin Exp Nephrol, 2012. **16**(1): p. 44-8.

253. Croissandeau, G., et al., *Increased stress-induced analgesia in mice lacking the proneuropeptide convertase PC2.* Neurosci Lett, 2006. **406**(1-2): p. 71-5.

254. Espinosa, V.P., et al., *Differential regulation of prohormone convertase 1/3, prohormone convertase 2 and phosphorylated cyclic-AMP-response element binding protein by short-term and long-term morphine treatment:*

implications for understanding the "switch" to opiate addiction. Neuroscience, 2008. **156**(3): p. 788-99.

255. Lenz, O., et al., *The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P.* Proc Natl Acad Sci U S A, 2001. **98**(22): p. 12701-5.

256. Hallenberger, S., et al., *Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gp160.* Nature, 1992. **360**(6402): p. 358-61.

257. Hatta, M., et al., *Molecular basis for high virulence of Hong Kong H5N1 influenza A viruses.* Science, 2001. **293**(5536): p. 1840-2.

258. Bastianelli, G., et al., *Computational reverse-engineering of a spider-venom derived peptide active against Plasmodium falciparum SUB1.* PLoS One, 2011. **6**(7): p. e21812.

259. Young, J.A. and R.J. Collier, *Anthrax toxin: receptor binding, internalization, pore formation, and translocation.* Annu Rev Biochem, 2007. **76**: p. 243-65.

260. Gordon, V.M., A. Rehemtulla, and S.H. Leppla, *A role for PACE4 in the proteolytic activation of anthrax toxin protective antigen.* Infect Immun, 1997. **65**(8): p. 3370-5.

261. Abrami, L., et al., *The pore-forming toxin proaerolysin is activated by furin.* J Biol Chem, 1998. **273**(49): p. 32656-61.

262. Steiner, D.F., *The proprotein convertases.* Curr Opin Chem Biol, 1998. **2**(1): p. 31-9.

263. Seidah, N.G. and A. Prat, *The biology and therapeutic targeting of the proprotein convertases.* Nat Rev Drug Discov, 2012. **11**(5): p. 367-83.

264. Selkoe, D.J., *Translating cell biology into therapeutic advances in Alzheimer's disease.* Nature, 1999. **399**(6738 Suppl): p. A23-31.

265. Checler, F., *Processing of the beta-amyloid precursor protein and its regulation in Alzheimer's disease.* J Neurochem, 1995. **65**(4): p. 1431-44.

266. Greenfield, J.P., et al., *Endoplasmic reticulum and trans-Golgi network generate distinct populations of Alzheimer beta-amyloid peptides.* Proc Natl Acad Sci U S A, 1999. **96**(2): p. 742-7.

267. van Wamelen, D.J., et al., *Decreased hypothalamic prohormone convertase expression in huntington disease patients.* J Neuropathol Exp Neurol, 2013. **72**(12): p. 1126-34.

268. Essalmani, R., et al., *In vivo functions of the proprotein convertase PC5/6 during mouse development: Gdf11 is a likely substrate.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5750-5.

269. Szumska, D., et al., *VACTERL/caudal regression/Currarino syndrome-like malformations in mice with mutation in the proprotein convertase Pcsk5.* Genes Dev, 2008. **22**(11): p. 1465-77.

270. Bassi, D.E., et al., *Elevated furin expression in aggressive human head and neck tumors and tumor cell lines.* Mol Carcinog, 2001. **31**(4): p. 224-32.

271. Turpeinen, H., et al., *Proprotein convertases in human atherosclerotic plaques: the overexpression of FURIN and its substrate cytokines BAFF and APRIL.* Atherosclerosis, 2011. **219**(2): p. 799-806.

272. Ornatowski, W., et al., *Elevated furin levels in human cystic fibrosis cells result in hypersusceptibility to exotoxin A-induced cytotoxicity.* J Clin Invest, 2007. **117**(11): p. 3489-97.

273. Ganesh, S.K., et al., *Loci influencing blood pressure identified using a cardiovascular gene-centric array.* Hum Mol Genet, 2013. **22**(8): p. 1663-78.

274. Furney, S.J., et al., *Genome-wide association with MRI atrophy measures as a quantitative trait locus for Alzheimer's disease.* Mol Psychiatry, 2011. **16**(11): p. 1130-8.

275. Ahmeti, K.B., et al., *Age of onset of amyotrophic lateral sclerosis is modulated by a locus on 1p34.1.* Neurobiol Aging, 2013. **34**(1): p. 357 e7-19.

276. Constam, D.B. and E.J. Robertson, *SPC4/PACE4 regulates a TGFbeta signaling network during axis formation.* Genes Dev, 2000. **14**(9): p. 1146-55.

277. Arning, L., et al., *VNTR Polymorphism Is Associated with Degree of Handedness but Not Direction of Handedness.* PLoS One, 2013. **8**(6): p. e67251.

278. Bassi, D.E., et al., *Enhanced aggressiveness of benzopyrene-induced squamous carcinomas in transgenic mice overexpressing the proprotein convertase PACE4 (PCSK6).* Mol Carcinog, 2014.

279. Longuespee, R., et al., *Implications of Proprotein Convertases in Ovarian Cancer Cell Proliferation and Tumor Progression: Insights for PACE4 as a Therapeutic Target.* Transl Oncol, 2014.

280. Malfait, A.M., et al., *A role for PACE4 in osteoarthritis pain: evidence from human genetic association and null mutant phenotype.* Ann Rheum Dis, 2012. **71**(6): p. 1042-8.

281. Tortorella, M.D., et al., *ADAMTS-4 (aggrecanase-1): N-terminal activation mechanisms.* Arch Biochem Biophys, 2005. **444**(1): p. 34-44.

282. Tsuji, A., et al., *Identification of novel cDNAs encoding human kexin-like protease, PACE4 isoforms.* Biochem Biophys Res Commun, 1994. **200**(2): p. 943-50.

283. Yang, Y., et al., *Solution structure of proinsulin: connecting domain flexibility and prohormone processing.* J Biol Chem, 2010. **285**(11): p. 7847-51.

284. Day, R., et al., *Distribution and regulation of the prohormone convertases PC1 and PC2 in the rat pituitary.* Mol Endocrinol, 1992. **6**(3): p. 485-97.

285. Ferraz-de-Souza, B. and J.C. Achermann, *Disorders of adrenal development.* Endocr Dev, 2008. **13**: p. 19-32.

286. Tomita, T., *Immunocytochemical localization of prohormone convertase 1/3 and 2 in gastrointestinal carcinoids.* Endocr Pathol, 2001. **12**(2): p. 137-45.

287. Jin, L., et al., *Distribution and regulation of proconvertases PC1 and PC2 in human pituitary adenomas.* Pituitary, 1999. **1**(3-4): p. 187-95.

288. Lankat-Buttgereit, B., et al., *Knockdown of Pdcd4 results in induction of proprotein convertase 1/3 and potent secretion of chromogranin A and secretogranin II in a neuroendocrine cell line.* Biol Cell, 2008. **100**(12): p. 703-15.

289. Cheng, M., et al., *Pro-protein convertase gene expression in human breast cancer.* Int J Cancer, 1997. **71**(6): p. 966-71.

290. Mbikay, M., et al., *Effects of rs6234/rs6235 and rs6232/rs6234/rs6235 PCSK1 single-nucleotide polymorphism clusters on proprotein convertase 1/3 biosynthesis and activity.* Mol Genet Metab, 2011. **104**(4): p. 682-7.

291. Strawbridge, R.J., et al., *Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes.* Diabetes, 2011. **60**(10): p. 2624-34.

292. Heni, M., et al., *Association of obesity risk SNPs in PCSK1 with insulin sensitivity and proinsulin conversion.* BMC Med Genet, 2010. **11**: p. 86.

293. Refaie, S., et al., *Disruption of proprotein convertase 1/3 (PC1/3) expression in mice causes innate immune defects and uncontrolled cytokine secretion.* J Biol Chem, 2012. **287**(18): p. 14703-17.

294. Posner, S.F., et al., *Stepwise posttranslational processing of progrowth hormone-releasing hormone (proGHRH) polypeptide by furin and PC1.* Endocrine, 2004. **23**(2-3): p. 199-213.

295. Berman, Y., et al., *Defective prodynorphin processing in mice lacking prohormone convertase PC2.* J Neurochem, 2000. **75**(4): p. 1763-70.

296. Furuta, M., et al., *Incomplete processing of proinsulin to insulin accompanied by elevation of Des-31,32 proinsulin intermediates in islets of mice lacking active PC2.* J Biol Chem, 1998. **273**(6): p. 3431-7.

297. Furuta, M., et al., *Defective prohormone processing and altered pancreatic islet morphology in mice lacking active SPC2.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6646-51.

298. Furuta, M., et al., *Severe defect in proglucagon processing in islet A-cells of prohormone convertase 2 null mice.* J Biol Chem, 2001. **276**(29): p. 27197-202.

299. Gyamera-Acheampong, C. and M. Mbikay, *Proprotein convertase subtilisin/kexin type 4 in mammalian fertility: a review.* Hum Reprod Update, 2009. **15**(2): p. 237-47.

300. Seidah, N.G., et al., *Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell-specific Kex2/subtilisin-like proprotein convertase.* Mol Endocrinol, 1992. **6**(10): p. 1559-70.

301. Nakayama, K., et al., *Identification of the fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease. Its testis-specific expression.* J Biol Chem, 1992. **267**(9): p. 5897-900.

302. Mbikay, M., et al., *Impaired fertility in mice deficient for the testicular germ-cell protease PC4.* Proc Natl Acad Sci U S A, 1997. **94**(13): p. 6842-6.

303. Oexle, K., et al., *Novel association to the proprotein convertase PCSK7 gene locus revealed by analysing soluble transferrin receptor (sTfR) levels.* Hum Mol Genet, 2011. **20**(5): p. 1042-7.

304. Lin, L., et al., *Soluble hemojuvelin is released by proprotein convertase-mediated cleavage at a conserved polybasic RNRR site.* Blood Cells Mol Dis, 2008. **40**(1): p. 122-31.

305. Guillemot, J., et al., *Implication of the proprotein convertases in iron homeostasis: proprotein convertase 7 sheds human transferrin receptor 1 and furin activates hepcidin.* Hepatology, 2013. **57**(6): p. 2514-24.

306. Besnard, J., et al., *Automated design of ligands to polypharmacological profiles.* Nature, 2012. **492**(7428): p. 215-20.

307. Weisz, O.A., *Acidification and protein traffic.* Int Rev Cytol, 2003. **226**: p. 259-319.

308. Sandvig, K., et al., *Acidification of the cytosol inhibits endocytosis from coated pits.* J Cell Biol, 1987. **105**(2): p. 679-89.

309. Sandvig, K., et al., *Inhibition of endocytosis from coated pits by acidification of the cytosol.* J Cell Biochem, 1988. **36**(1): p. 73-81.

310. Cosson, P., et al., *Low cytoplasmic pH inhibits endocytosis and transport from the trans-Golgi network to the cell surface.* J Cell Biol, 1989. **108**(2): p. 377-87.

311. Samuelson, A.C., et al., *Influence of cytosolic pH on receptor-mediated endocytosis of asialoorosomucoid.* Am J Physiol, 1988. **254**(6 Pt 1): p. C829-38.

312. Cardone, R.A., V. Casavola, and S.J. Reshkin, *The role of disturbed pH dynamics and the Na+/H+ exchanger in metastasis.* Nat Rev Cancer, 2005. **5**(10): p. 786-95.

313. Harguindey, S., et al., *Growth and trophic factors, pH and the Na+/H+ exchanger in Alzheimer's disease, other neurodegenerative diseases and cancer: new therapeutic possibilities and potential dangers.* Curr Alzheimer Res, 2007. **4**(1): p. 53-65.

314. Syntichaki, P., C. Samara, and N. Tavernarakis, *The vacuolar H+ - ATPase mediates intracellular acidification required for neurodegeneration in C. elegans.* Curr Biol, 2005. **15**(13): p. 1249-54.

315. Brett, C.L., et al., *Genome-wide analysis reveals the vacuolar pH-stat of Saccharomyces cerevisiae.* PLoS One, 2011. **6**(3): p. e17619.

316. Renstrom, E., R. Ivarsson, and S.B. Shears, *Inositol 3,4,5,6-tetrakisphosphate inhibits insulin granule acidification and fusogenic potential.* J Biol Chem, 2002. **277**(30): p. 26717-20.

317. Oliver, K.G., et al., *The interconversion of inositol 1,3,4,5,6-pentakisphosphate and inositol tetrakisphosphates in AR4-2J cells.* J Biol Chem, 1992. **267**(30): p. 21528-34.

318. Maldonado, E.N. and J.J. Lemasters, *ATP/ADP ratio, the missed connection between mitochondria and the Warburg effect.* Mitochondrion, 2014. **19 Pt A**: p. 78-84.

319. Mochly-Rosen, D., K. Das, and K.V. Grimes, *Protein kinase C, an elusive therapeutic target?* Nat Rev Drug Discov, 2012. **11**(12): p. 937-57.

320. Seksek, O., J. Biwersi, and A.S. Verkman, *Direct measurement of trans-Golgi pH in living cells and regulation by second messengers.* J Biol Chem, 1995. **270**(10): p. 4967-70.

321. Pace, C.S., *Role of pH as a transduction device in triggering electrical and secretory responses in islet B cells.* Fed Proc, 1984. **43**(9): p. 2379-84.

322. Smith, J.S. and C.S. Pace, *Modification of glucose-induced insulin release by alteration of pH.* Diabetes, 1983. **32**(1): p. 61-6.

323. Best, L., et al., *Is intracellular pH a coupling factor in nutrient-stimulated pancreatic islets?* J Mol Endocrinol, 1988. **1**(1): p. 33-8.

324. Gunawardana, S.C., et al., *Nutrient-stimulated insulin secretion in mouse islets is critically dependent on intracellular pH.* BMC Endocr Disord, 2004. **4**(1): p. 1.

325. Henomatsu, N., et al., *Inhibition of intracellular transport of newly synthesized prolactin by bafilomycin A1 in a pituitary tumor cell line, GH3 cells.* Eur J Cell Biol, 1993. **62**(1): p. 127-39.

326. Grimsley, G.R., J.M. Scholtz, and C.N. Pace, *A summary of the measured pK values of the ionizable groups in folded proteins.* Protein Sci, 2009. **18**(1): p. 247-51.

327. Di Russo, N.V., et al., *pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4.* PLoS Comput Biol, 2012. **8**(11): p. e1002761.

328. Kukic, P., et al., *Improving the analysis of NMR spectra tracking pH-induced conformational changes: removing artefacts of the electric field on the NMR chemical shift.* Proteins, 2010. **78**(4): p. 971-84.

329. Tanford, C., *Ionization-linked Changes in Protein Conformation. I. Theory.* J Am Chem Soc, 1961. **83**(7): p. 1628-1634.

330. Wyman, J., Jr., *Heme proteins.* Adv Protein Chem, 1948. **4**: p. 407-531.

331. Flegal, K.M., et al., *Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010.* JAMA, 2012. **307**(5): p. 491-7.

332. Creemers, J.W., R.S. Jackson, and J.C. Hutton, *Molecular and cellular regulation of prohormone processing.* Semin Cell Dev Biol, 1998. **9**(1): p. 3-10.

333. Choquet, H., et al., *Contribution of common PCSK1 genetic variants to obesity in 8,359 subjects from multi-ethnic American population.* PLoS One, 2013. **8**(2): p. e57857.

334. Chan, P. and J. Warwicker, *Evidence for the adaptation of protein pH-dependence to subcellular pH.* BMC Biol, 2009. **7**: p. 69.

335. Lynch, M. and J.S. Conery, *The evolutionary fate and consequences of duplicate genes.* Science, 2000. **290**(5494): p. 1151-5.

336. Feng, D.F., G. Cho, and R.F. Doolittle, *Determining divergence times with a protein clock: update and reevaluation.* Proc Natl Acad Sci U S A, 1997. **94**(24): p. 13028-33.

337. Izaguirre, G., et al., *Identification of serpin determinants of specificity and selectivity for furin inhibition through studies of alpha1PDX (alpha1-protease inhibitor Portland)-serpin B8 and furin active-site loop chimeras.* J Biol Chem, 2013. **288**(30): p. 21802-14.

338. Cordelier, P., M.A. Zern, and D.S. Strayer, *HIV-1 proprotein processing as a target for gene therapy.* Gene Ther, 2003. **10**(6): p. 467-77.

339. Lei, X., et al., *Hepatic overexpression of the prodomain of furin lessens progression of atherosclerosis and reduces vascular remodeling in response to injury.* Atherosclerosis, 2014. **236**(1): p. 121-30.

340. Seidah, N.G., *Proprotein convertase subtilisin kexin 9 (PCSK9) inhibitors in the treatment of hypercholesterolemia and other pathologies.* Curr Pharm Des, 2013. **19**(17): p. 3161-72.

341. Becker, G.L., et al., *Highly potent inhibitors of proprotein convertase furin as potential drugs for treatment of infectious diseases.* J Biol Chem, 2012. **287**(26): p. 21992-2003.

342. Wang, L., et al., *Prodomain mutations at the subtilisin interface: correlation of binding energy and the rate of catalyzed folding.* Biochemistry, 1995. **34**(47): p. 15415-20.

343. Cash, P.W., et al., *Synthesis of the pro-peptide of subtilisin BPN'.* Pept Res, 1989. **2**(4): p. 292-6.

344. Liu, Y. and B. Kuhlman, *RosettaDesign server for protein design.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W235-8.

345. O'Rahilly, S., et al., *Brief report: impaired processing of prohormones associated with abnormalities of glucose homeostasis and adrenal function.* N Engl J Med, 1995. **333**(21): p. 1386-90.

346. Bandsma, R.H., et al., *From diarrhea to obesity in prohormone convertase 1/3 deficiency: age-dependent clinical, pathologic, and enteroendocrine characteristics.* J Clin Gastroenterol, 2013. **47**(10): p. 834-43.

347. Frank, G.R., et al., *Severe obesity and diabetes insipidus in a patient with PCSK1 deficiency.* Mol Genet Metab, 2013. **110**(1-2): p. 191-4.

348. Manning, A.K., et al., *A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance.* Nat Genet, 2012. **44**(6): p. 659-69.

349. Qi, Q., et al., *Association of PCSK1 rs6234 with obesity and related traits in a Chinese Han population.* PLoS One, 2010. **5**(5): p. e10590.

350. Chang, Y.C., et al., *Common PCSK1 haplotypes are associated with obesity in the Chinese population.* Obesity (Silver Spring), 2010. **18**(7): p. 1404-9.

351. Ahlqvist, E., et al., *A common variant upstream of the PAX6 gene influences islet function in man.* Diabetologia, 2012. **55**(1): p. 94-104.

352. Anini, Y., et al., *Genetic deficiency for proprotein convertase subtilisin/kexin type 2 in mice is associated with decreased adiposity and protection from dietary fat-induced body weight gain.* Int J Obes (Lond), 2010. **34**(11): p. 1599-607.

353. Miller, R., et al., *Selective roles for the PC2 processing enzyme in the regulation of peptide neurotransmitter levels in brain and peripheral neuroendocrine tissues of PC2 deficient mice.* Neuropeptides, 2003. **37**(3): p. 140-8.

354. Miller, R., et al., *Obliteration of alpha-melanocyte-stimulating hormone derived from POMC in pituitary and brains of PC2-deficient mice.* J Neurochem, 2003. **86**(3): p. 556-63.

355. Pan, H., et al., *The role of prohormone convertase-2 in hypothalamic neuropeptide processing: a quantitative neuropeptidomic study.* J Neurochem, 2006. **98**(6): p. 1763-77.

356. Rehfeld, J.F., I. Lindberg, and L. Friis-Hansen, *Increased synthesis but decreased processing of neuronal proCCK in prohormone convertase 2 and 7B2 knockout animals.* J Neurochem, 2002. **83**(6): p. 1329-37.

357. Allen, R.G., et al., *Altered processing of pro-orphanin FQ/nociceptin and pro-opiomelanocortin-derived peptides in the brains of mice expressing defective prohormone convertase 2.* J Neurosci, 2001. **21**(16): p. 5864-70.

358. Yoshida, H., et al., *Association of the prohormone convertase 2 gene (PCSK2) on chromosome 20 with NIDDM in Japanese subjects.* Diabetes, 1995. **44**(4): p. 389-93.

359. Jonsson, A., et al., *Effect of a common variant of the PCSK2 gene on reduced insulin secretion.* Diabetologia, 2012. **55**(12): p. 3245-51.

360. Leak, T.S., et al., *Association of the proprotein convertase subtilisin/kexin-type 2 (PCSK2) gene with type 2 diabetes in an African American population.* Mol Genet Metab, 2007. **92**(1-2): p. 145-50.

361. Zheng, X., et al., *Association of type 2 diabetes susceptibility genes (TCF7L2, SLC30A8, PCSK1 and PCSK2) and proinsulin conversion in a Chinese population.* Mol Biol Rep, 2012. **39**(1): p. 17-23.

362. Yoshida, T., et al., *Association of gene polymorphisms with chronic kidney disease in Japanese individuals.* Int J Mol Med, 2009. **24**(4): p. 539-47.

363. Murea, M., et al., *Genome-wide association scan for survival on dialysis in African-Americans with type 2 diabetes.* Am J Nephrol, 2011. **33**(6): p. 502-9.

364. O'Donnell, C.J., et al., *Genome-wide association study for subclinical atherosclerosis in major arterial territories in the NHLBI's Framingham Heart Study.* BMC Med Genet, 2007. **8 Suppl 1**: p. S4.

365. Fujimaki, T., et al., *Association of genetic variants in SEMA3F, CLEC16A, LAMA3, and PCSK2 with myocardial infarction in Japanese individuals.* Atherosclerosis, 2010. **210**(2): p. 468-73.

366. Lei, R.X., et al., *Influence of a single nucleotide polymorphism in the P1 promoter of the furin gene on transcription activity and hepatitis B virus infection.* Hepatology, 2009. **50**(3): p. 763-71.

367. Qiu, Q., et al., *Role of pro-IGF-II processing by proprotein convertase 4 in human placental development.* Proc Natl Acad Sci U S A, 2005. **102**(31): p. 11047-52.

368. Iatan, I., et al., *Genetic variation at the proprotein convertase subtilisin/kexin type 5 gene modulates high-density lipoprotein cholesterol levels.* Circ Cardiovasc Genet, 2009. **2**(5): p. 467-75.

369. Li, J.P., et al., *The association between paired basic amino acid cleaving enzyme 4 gene haplotype and diastolic blood pressure.* Chin Med J (Engl), 2004. **117**(3): p. 382-8.

370. Middelberg, R.P., et al., *Genetic variants in LPL, OASL and TOMM40/APOE-C1-C2-C4 genes are associated with multiple cardiovascular-related traits.* BMC Med Genet, 2011. **12**: p. 123.