Furilisin: A Variant of Subtilisin BPN′ Engineered for Cleaving Tribasic Substrates†

Marcus D. Ballinger,‡ Jeffrey Tom,§ and James A. Wells*,‡

Departments of Protein Engineering and Bioorganic Chemistry, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080

Received June 26, 1996⊗

ABSTRACT: The serine protease, subtilisin BPN′, was engineered to cleave proteins after tribasic sequences in a manner that resembles the substrate specificity of furin, one of the mammalian subtilisin homologs that processes prohormones. As a starting point we used a double mutant of subtilisin BPN′ (N62D/G166D) that showed substantial preference for cleaving after sequences having consecutive dibasic residues (namely, at the P1 and P2 substrate positions) [Ballinger *et al.* (1995) *Biochemistry 34*, 13312−13319]. Additional specificity for basic residues was engineered at the P4 position by introducing subtilisin-to-furin substitutions at three hydrophobic residues that composed the S4 subsite (Y104, I107, and L126). Initial attempts to incorporate a Y104D or I107E mutation or the Y104D/I107E double mutation into the dibasic specific enzyme failed to generate the processed enzyme. The problem was traced to the inability of the mutant prosubtilisins to process themselves and fold correctly. Replacing the natural processing site sequence (AHAY) with a good furin substrate sequence (RHKR) resulted in expression of the triple subtilisin mutant (N62D/Y104D/G166D) we call "furilisin". Furilisin hydrolyzes synthetic tribasic substrates (succinyl-RAKR-pNA or succinyl-KAKR-pNA) with high catalytic efficiency ($k_{cat}/K_m > 3 \times 10^5$ M^{-1} s^{-1}) and discriminates in favor of Arg versus Ala at the P4 position by a factor of 360. The overall specificity change versus the wild-type enzyme was dramatic. For example, succinyl-RAKR-pNA was cleaved ∼60 000 times faster than succinyl-AAPF-pNA, a good substrate for wild-type subtilisin. Similarly, furilisin was inhibited ($K_i^* = 29$ nM) by a variant of the turkey ovomucoid third domain inhibitor that contained an engineered furin substrate site (RCKR↓) [Lu et al. (1993) *J. Biol. Chem. 268*, 14583−14585] and not by one having a good wild-type subtilisin substrate sequence (ACTL↓). Interestingly, the extreme changes in substrate specificity resulted from substantial synergy between the engineered subsites. These studies provide a basic example of how to manipulate substrate specificity in a modular fashion, thereby creating an engineered enzyme that may be useful as a protein processing tool.

The use of expression systems in which proteins are synthesized as fusions to affinity domains has increased dramatically in recent years, owing to the strong advantage gained in affinity purification (Carter, 1990; Uhlen & Moks, 1990; Nilsson et al., 1992; Beck et al., 1994; LaVallie & McCoy, 1995). An essential reagent for this strategy is an efficient and specific protease for cleavage and release of the purified protein of interest. Ideally, the site-specific protease should have highly restrictive substrate specificity, high catalytic power, and stability and should be available in large amounts.

Few naturally occurring proteases exhibit all of these properties. For instance, the growing family of mammalian subtilisin-like prohormone convertases are exquisitely specific proteases [for reviews see Steiner et al. (1992) and Smeekens (1993)]. These enzymes, for which furin is an example, have substrate sequence requirements for tribasic residues (Bresnahan et al., 1990; Wise et al., 1990; Hosaka et al., 1991; Matthews et al., 1994). However, the prohormone convertases are membrane associated, and soluble forms are poorly expressed in recombinant systems (Bravo et al., 1994; Matthews et al., 1994). At the other end of the spectrum, scavenging enzymes such as bacterial subtilisins can be expressed very well and in secreted form (Wells et al., 1983) but have a broad substrate specificity (Kraut, 1977; Grøn et al., 1992).

Previously we engineered a variant of subtilisin BPN′ [1] with the dibasic substrate specificity (Ballinger et al., 1995) that is similar to the a yeast subtilisin homolog, kex2 (Mizuno et al., 1988). This subtilisin BPN′ variant, called kexilisin, contained acidic mutations in the S1 and S2 subsites that were designed from considerations of the X-ray structure of subtilisin and sequence alignments with kex2 and other prohormone convertases. Point mutations at the base of the substrate binding pockets yielded substantial specificity

† This work was supported in part by an NRSA Fellowship (M.D.B.) from NIH Postdoctoral Training Grant GM16549−01.

* To whom correspondence should be addressed.

‡ Department of Protein Engineering, Genentech, Inc.

§ Department of Bioorganic Chemistry, Genentech, Inc.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1996.

[1] Abbreviations: subtilisin BPN′, subtilisin from *Bacillus amyloliquefaciens*; kexilisin, N62D/G166D subtilisin BPN′; furilisin, N62D/Y104D/G166D subtilisin BPN′; pNA, *p*-nitroanilide; OMTKY3, turkey ovomucoid third domain; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; vWF, von Willebrand factor. Mutants are designated by the single letter amino acid code for the wild-type residue followed by their position and replacement residue. Multiple mutants are indicated by component single mutants separated by slashes. Substrates are described in triplet or single letter code as P*n*···P2−P1−P1′-P2′···P*n*′, where the scissile bond is between the P1 and P1′ residues (Schechter & Berger, 1967). For substrate (or inhibitor) sequences, ↓ denotes the scissile bond (or the bond following P1).

S0006-2960(96)01543-7 CCC: $12.00

changes and preserved the high activity and expression levels of the native subtilisin BPN′.

To further restrict the specificity, we sought to add the requirement for a P4 basic residue as found in furin. Several studies have shown that it is possible to modulate the substrate preference at the P4 position by mutations at Y104, I107, and L126 or their equivalents [Bech et al., 1992, 1993; Teplyakov et al., 1992; Rheinnecker et al., 1993, 1994; Sørensen et al., 1993; reviewed in Perona and Craik (1995)]. These studies focused primarily on altering the size preference for neutral residues. In contrast to the hydrophobic nature of the S4 subsite in subtilisin BPN′, functional studies (Creemers et al., 1993) and homology modeling (Siezen et al., 1994; Lipkind et al., 1995) of furin and other prohormone convertases highlight several negatively charged residues in the S4 subsite that could potentially direct basic specificity.

Subtilisins are initially expressed as preproenzymes (Power et al., 1986), and the propeptide acts as an intramolecular chaperone that is critical for folding of the mature protease domain (Shinde & Inoue, 1994). As revealed by the recent crystal structure of the subtilisin−propeptide complex (Bryan et al., 1995; Gallagher et al., 1995), the pro-mature processing site is positioned in the substrate binding cleft, allowing for intramolecular cleavage by the enzyme upon folding. An intramolecular cleavage mechanism was also strongly supported by experiments of Shinde and Inoue in which an active site His64Ala variant of subtilisin E, which operates on HisP2 substrates by a substrate-assisted catalysis mechanism (Carter & Wells, 1987), could mature itself only upon introducing a histidine at the P2 position of the processing site (Shinde & Inoue, 1995). Kex2 and other prohormone convertases also process themselves via intramolecular cleavage (Siezen et al., 1995).

Here, we have engineered a multiple mutant of subtilisin BPN′ that is specific for tribasic sequences, called "furilisin". This was accomplished by combining an acidic mutation at the S4 site with the kexilisin S1 and S2 site mutations. The specificity of furilisin was so dramatically altered compared to wild-type subtilisin or kexilisin that we had to modify the processing site to contain a tribasic sequence to obtain expression. Interestingly, the cumulative effect of incorporating acidic substitutions had a synergistic effect on the preference for basic residues. This may be a general effect when engineering specificity for charged substrates in enzymes and proteins.

MATERIALS AND METHODS

Materials. Oligonucleotides and *p*-nitroanilide substrates were synthesized and purified by the Bioorganic Chemistry Department at Genentech. Chromatography columns were from Pharmacia Biotechnology. Turkey ovomucoid third domain inhibitors were a generous gift of Dr. Michael Laskowski in the Chemistry Department at Purdue University.

Construction and Purification of Subtilisin Mutants. Site-directed mutations were introduced into kexilisin, N62D/G166D (Ballinger et al., 1995), or kexilisin containing additional changes in the processing site and S4 subsite, A(−4)R/A(−2)K/Y(−1)R/N62D/Y104D/G166D. The mutagenesis of the subtilisin BPN′ gene was performed in the phagemid pSS5 (Carter & Wells, 1987; Ballinger et al., 1995) using the method of Kunkel et al. (1991). Oligonucleotides used for mutagenesis were designed with overlaps of 12 base pairs on either side of mutations. For processing site and S4 subsite combination mutants, both oligonucleotides were simultaneously introduced in the same mutagenesis procedure. Mutations were confirmed by dideoxy sequencing (Sanger et al., 1977), and double-stranded mutant DNA (produced from either 294MM or JM101 *Escherichia coli*) was transformed into a protease-deficient strain (BG2036) of *Bacillus subtilis* (Yang et al., 1984). Transformants were plated out on skim milk plates.

Subtilisin variants were cultured and purified by ethanol precipitation followed by mono-S cation-exchange chromatography as described previously (Ballinger et al., 1995). Purified wild-type and mutant subtilisins were >95% pure as determined by SDS−PAGE.

Kinetic Analyses. Subtilisins were assayed by the initial rates of hydrolysis of pNA tetrapeptide substrates in 0.4 mL of 20 mM Tris-HCl (pH 8.2), 4% (v/v) dimethyl sulfoxide at 25 ± 0.2 °C as described previously (Estell et al., 1986). Kinetic measurements for obtaining Michaelis−Menten plots were made using a Kontron Uvikon 860 spectrophotometer. Enzyme concentrations were determined spectrophotometrically using $\epsilon_{280}^{0.1\%} = 1.17$ (Matsubara et al., 1965) or 1.12 for the Y104D-containing mutant, considering the calculated effect of Tyr replacement on the 280 nm absorbance (Freifelder, 1982). Initial rates were determined for 9−12 different substrate concentrations over the range of 0.002−1.5 mM that bracketed K_m. Substrate concentrations were determined by allowing pNA hydrolysis to proceed to completion, often assisted by the addition of 0.05 M (final) NaOH, and measuring the absorbance at 410 nm ($\epsilon_{410} = 8480$ cm^{-1} M^{-1}). For several poor enzyme−substrate combinations, in which pNA hydrolysis did not proceed to completion enzymatically, partially digested reaction products were separated by reverse-phase HPLC on a Vydac C18 column and characterized by mass spectrometry. The furilisin digestions of succinyl-RGAR-pNA and succinyl-RGKF-pNA yielded primarily the products of *p*-nitroanilide hydrolysis; however, some product (20−30% of the total) resulted from cleavage at succinyl-RGA↓R-pNA or succinyl-RGK↓F-pNA, respectively. For kexilisin digestions of succinyl-RGKR-pNA, cleavage was almost exclusively at the succinyl-RGK↓R-pNA, making kinetic parameters for succinyl-RGKR↓pNA unobtainable. V_{max} and K_m values were obtained from initial rate versus substrate concentration data by nonlinearly least-squares fitting to the Michaelis−Menten equation using the program KaleidaGraph (Synergy Software, Reading, PA).

Inhibition by turkey ovomucoid third domains was measured using the methods for tight-binding inhibitors of Empie and Laskowski (1982), in which the enzyme (1 nM for wild-type subtilisin, 3 nM for furilisin) and various concentrations of inhibitor were incubated at 21 °C for 20 h in 0.005% Triton X-100 and 20 mM Tris-HCl (pH 8.2) in 96-well polystyrene plates. Residual activity toward 75 μM succinyl-AAPF-pNA (wild-type subtilisin) or 45 μM succinyl-RGKR-pNA (furilisin), added at $^1/_{20}$th final volume in 20% DMSO, was measured using an SLT microplate reader using Deltasoft II software. Apparent equilibrium dissociation values (K_i*) were calculated by fitting the data to the equation (Morrison, 1969; Bieth, 1974; Williams & Morrison, 1979):

$$\frac{V_i}{V_0} = 1 - \frac{[E_0] + [I_0] + K_i^* - \sqrt{([E_0] + [I_0] + K_i^*)2 - (4[E_0][I_0])}}{2[E_0]} \quad (1)$$

where V_i/V_0 is the fractional activity (initial inhibited rate divided by the initial uninhibited rate), $[E_0]$ is the total enzyme concentration, and $[I_0]$ is the total inhibitor concentration. The substrate concentrations ($0.5K_m$) were minimized to avoid substrate-induced dissociation during the activity assay (Empie & Laskowski, 1982) but maintained at a level that would still yield enough activity to allow accurate rate measurements.

Furilisin−RAKR Complex Model. The Arg-Ala-Lys-Arg peptide sequence and furilisin mutations were modeled into the structure of Ala-Ala-Pro-boroPhe inhibitor complexed with subtilisin BPN′ (Butcher and Kossiakoff, unpublished results) using Insight II software. The main-chain atoms and unchanged residues were held constant except for a slight change in the P3Ala, P4Arg backbone positions due to the P2 Pro to Lys substitution. Hydrogen bonds from P1Arg to Asp166, P2Lys to Asp62, and P4Arg to Asp104 were modeled by manipulating side chains while maintaining torsion angles in anti or gauche conformations (±20°) and avoiding unfavorable van der Waals interactions with the protein. Installing the ArgP1 and Asp166 side chain required displacement of two ordered water molecules from the S1 subsite.

RESULTS

S4 Specificity Design. The X-ray crystal structure of subtilisin BPN′ complexed with a succinyl-Ala-Ala-Pro-boroPhe inhibitor (Butcher and Kossiakoff, unpublished results) reveals three residues whose side chains contribute to forming the S4 subsite. Tyr104 forms the outside edge of the cavity, Ile107 lies at the base of the pocket, and Leu126 forms a edge opposite Tyr104 (Figure 1A). These residues have been mutated to other hydrophobic residues (Rheinnecker et al., 1993, 1994) and caused shifts in the preference toward large hydrophobic substrates. In subtilisin 309 (another *Bacillus* subtilisin), Arg and Asp substitutions (among others) were made at position 104 (Bech et al., 1992, 1993; Sørensen et al., 1993). However, these mutants had low activity, even versus complementary charged substrates (Bech et al., 1992).

The eukaryotic subtilisin homologs (furin, PC2, and PC3) show a strong consensus at each of the S4 subsite positions compared to subtilisin BPN′ (Table 1). These enzymes have Asp at position 104, except for kex2 which has Thr and is not P4 basic specific; Glu107 and Trp126 are exclusively conserved. We installed furin-like substitutions into kexilisin either singly (for Y104D and I107E) or in combinations. However, none of these variants expressed mature protease, as seen by the lack of both tetrapeptide hydrolyzing activity or mature subtilisin gel bands in cell culture supernatants (see below). Interestingly, the mature subtilisin product from these mutants was not obtained even when coculturing with A48E "helper" subtilisin (data not shown), a strategy used to express catalytic triad knockout mutants of subtilisin (Carter & Wells, 1987, 1988). This suggests that good substrate−subsite complementarity is required for folding the protease.

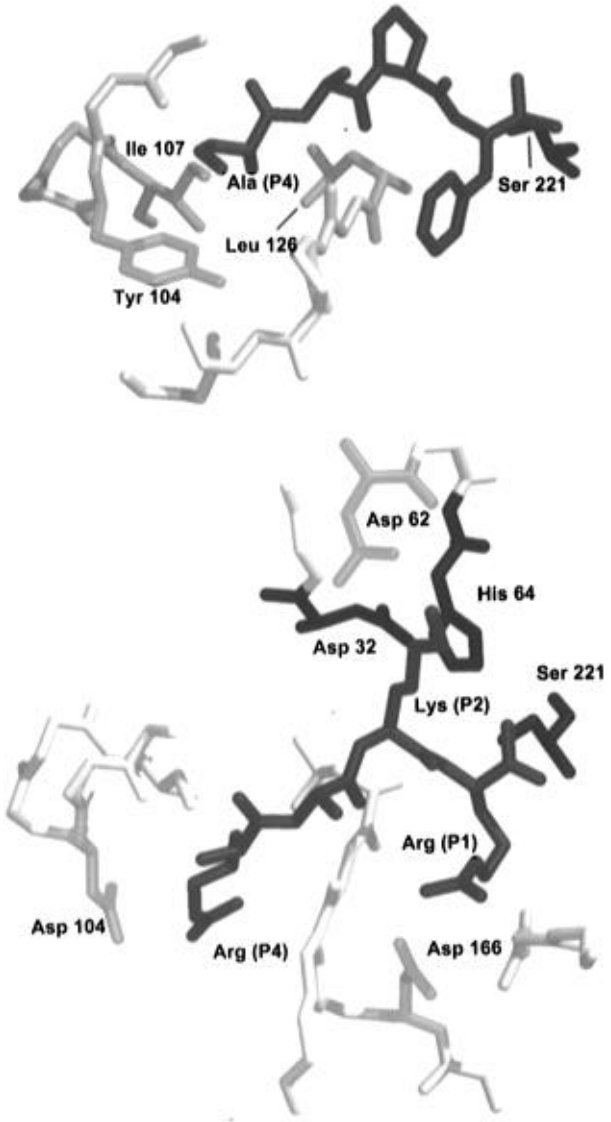

FIGURE 1: (A, top) Structure of an inhibitor, succinyl-Ala-Ala-Pro-boroPhe inhibitor (red), bound to the active site of subtilisin BPN′ (K. Butcher and T. Kossiakoff, unpublished results). The free carboxylate and methylene groups of the succinyl group are disordered and not displayed. Residues in the S4 binding site are shown in yellow, and Ser221 (bound to the boronate) is in blue. (B, bottom) A model for an Arg-Ala-Lys-Arg peptide bound to furilisin based on the structure in (A). Binding site mutations are shown in yellow and catalytic triad residues in blue.

Table 1: Sequence Alignments for Portions of the S4 Site of Subtilisins[a]

	S4 site	
	104 107	126
Subtilisin	GSGQYSWIING	NMSLGGP
Kex2	GDITTEDEAAS	SCSWGPA
Furin	GEVTDAVEARS	SASWGPE
PC2	PFMTDIIEASS	SASWGPT
PC3	GIVTDAIEASS	SASWGPN

[a] Numbering is based on the subtilisin BPN′ sequence [from Siezen et al. (1991)].

We reasoned that the cumulative active site substitutions could have caused too great an overall shift in the substrate specificity and thus prevented processing. Therefore, the wild-type autoprocessing site (Ala-His-Ala-Tyr↓Ala) was mutated to the furin-like substrate sequence Arg-His-Lys-

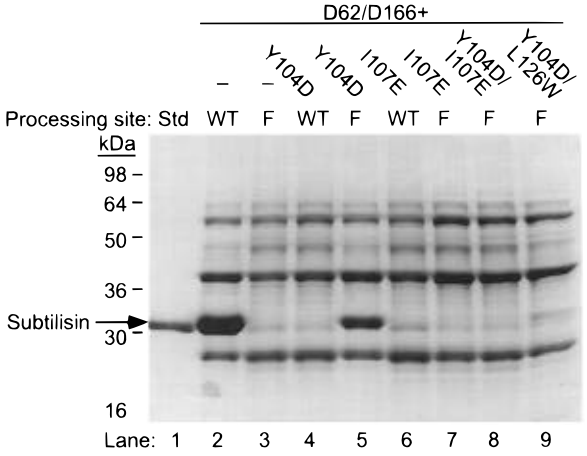

FIGURE 2: 14% SDS−PAGE of crude extracellular medium from subtilisin variants, showing the mutations in addition to N62D/G166D (kexilisin) background and the processing site. Processing sites were either AHAY↓A (WT) or RHKR↓A (F). The band running slightly below the subtilisin marker does not correspond to active subtilisin (as determined by spectrophotometric assay of the mixture vs AAPF-pNA or RAKR-pNA).

Arg↓Ala in combination with the 104 and 107 acidic mutations. As shown in Figure 2, the wild-type processing site (AHAY) is accepted sufficiently by kexilisin to allow processing and expression (lane 2). However, with the tribasic processing site (RHKR), kexilisin did not express (lane 3), presumably due to the inability of the wild-type S4 subsite to accommodate the P4Arg. In contrast, the processing site change allowed kexilisin containing the Y104D mutant (furilisin) to express (lane 5) and produced roughly 15 mg of enzyme/liter of culture. This suggested that the Y104D mutation induced a strong preference for P4 Arg over Ala. Variants containing the I107E mutation, however, could not be expressed even with the change in the processing site (lanes 6 and 7) or when combined with the Y104D substitution (lane 8). This could be due to the fact that I107 is more interior to the hydrophobic pocket and thus the introduction of the charged group imparts instability to the enzyme. In addition, the S4 subsite in subtilisin BPN′ may lack sufficient room to accommodate the Glu107 and ArgP4 interaction. The L126W mutation was introduced into kexilisin containing the Y104D mutation plus the RHKR↓ processing site to see if it would enhance the expression of the enzyme or its specificity toward tribasic substrates. Unfortunately, this additional substitution appeared detrimental to expression (lane 9); furthermore, assays of the crude supernatant indicated a substantially reduced activity toward tribasic substrates (data not shown).

Kinetic Analysis of N62D/Y104D/G166D Subtilisin ("Furilisin"). The mature N62D/Y104D/G166D subtilisin, which we call furilisin, was purified and analyzed along with kexilisin and wild-type subtilisin for their abilities to hydrolyze several tetrapeptide-pNA substrates (Table 2). The wild-type enzyme hydrolyzes the hydrophobic substrate, succinyl-AAPF-pNA, 140 or 30 300 times faster than the kexilisin or furilisin, respectively (Figure 3). Kexilisin hydrolyzes the dibasic substrate, succinyl-AAKR-pNA, 46 times faster than furilisin and 220 times faster than wild type. Furilisin hydrolyzes the tribasic substrates (succinyl-RAKR-pNA and succinyl-KAKR-pNA) with high catalytic efficiency ($>10^5$ M^{-1} s^{-1}); in contrast, cleavage of these substrates by wild-type subtilisin or kexilisin is so poor that assaying them was impossible due to predominant cleavage between Lys and Arg-pNA. This was found to be true even when the substrate series was changed such that Gly replaced Ala at the P3 position; this should have deterred cleavage between Lys and Arg-pNA because Gly is substantially worse than Ala as a P2 residue (Grøn et al., 1992). The alternate cleavage could be induced by the placement of Arg at P3, which is a favorable residue at this position (Grøn et al., 1992). Remarkably, both kexilisin and furilisin hydrolyzed their corresponding best di- or tribasic substrates with the same or higher catalytic efficiency (k_{cat}/K_m) as wild-type subtilisin with its favored substrate, succinyl-AAPF-pNA.

The preferences at the S4 subsites of kexilisin and furilisin were further explored using a succinyl-XGKR-pNA substrate series (Table 2, Figure 4). Kexilisin hydrolyzes all three hydrophobic substrates efficiently, with the order of preference Leu > Phe > Ala. Although data for the ArgP4 substrate could not be obtained due to incorrect cleavage, Grøn et al. (1992) showed that, for an alternate series, subtilisin prefers Ala over Arg by >200-fold in k_{cat}/K_m at this position. Furilisin reverses this preference such that Arg is favored over Ala by a factor of 360. The cleavage efficiency with Leu at P4 was also dramatically reduced; however, Phe was hydrolyzed with efficiency virtually equal to Arg. This suggests that the S4 subsite in furilisin can accommodate aromatic as well as basic residues.

The three furilisin mutations combine to yield a 60 000-fold discrimination in favor of succinyl-RAKR-pNA over succinyl-AAPF-pNA. This value is higher than would be predicted from combining the kexilisin dibasic (P1, P2) and furilisin P4 specificities in an additive fashion (i.e., the S2−S1 mutations together cause a 200-fold discrimination in favor of KR over PF; the S4 mutation causes a 60-fold favoring of R over A; the product is 12 000-fold). To further evaluate this effect, we analyzed the relative catalytic efficiencies of Arg and PheP1 substrates as a function of the substrate−subsite background. As shown in Figure 5, the wild-type enzyme favors Phe over Arg at P1 in the succinyl-AAPX-pNA substrate series by a factor of ∼20. The single G166D mutation in the S1 subsite causes a shift to nearly equal efficiency of hydrolysis. Changing the background to kexilisin with succinyl-AAKX-pNA substrates results in a significant preference for Arg over Phe at the P1 site (12-fold). Finally, the furilisin mutant with succinyl-RGKX-pNA enhances the P1 preference, up to 160-fold. In the latter series, the succinyl-RGKF-pNA kinetics are complicated slightly because ∼20% of the substrate is cleaved between succinyl-RGK and F-pNA as verified by HPLC analysis of reaction mixtures. However, this should only marginally affect the observed succinyl-RGKF↓pNA cleavage rates. For the S2−P2 interaction in furilisin, the cumulative effect was not as dramatic, although the Lys was preferred over Ala by a factor of 24.

Inhibition by Turkey Ovomucoid Third Domains. To further evaluate the specificity change in furilisin, we examined inhibition by two variants of turkey ovomucoid third domain (OMTKY3) mutants, a small potent inhibitor of many serine proteases (Empie & Laskowski, 1982). Residues 15−18 of the inhibitor occupy positions P4−P1 in the protease active site and have the sequence ACTL. The wild-type OMTKY3 (unavailable) and an R21M (P3′) variant are good inhibitors of bacterial subtilisins whereas the A15R/T17K/L18R variant is a moderate inhibitor (K_i = 91 nM)[2] of human furin (Lu et al., 1993). Wild-type subtilisin was

[2] Calculated from the reported K_a of 1.1×10^7 M (Lu et al., 1993) assuming $K_i = 1/K_a$.

Table 2: Kinetic Analysis of Wild-Type Subtilisin BPN′, Kexilisin, and Furilisin Mutants versus Succinyl-tetrapeptide-pNA Substrates[a]

	subtilisin mutant											
	WT			G166D			kexilisin (N62D/G166D)			furilisin (N62D/Y104D/G166D)		
succinyl-tetrapeptide substrate sequence	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})
AAPF	29	110	2.6×10^5	6.1	1000	5.8×10^3	3.4	1800	1.9×10^3	ND[b]	ND[b]	8.0×10^{0} [b]
AAPR	8.9	630	1.4×10^4	2.7	500	5.5×10^3	ND	ND	ND	ND	ND	ND
AAKR	2.8	1700	1.7×10^3 [c]	ND	ND	ND	15	41	3.7×10^5	0.27	34	8.1×10^3
AAKF	13	79	1.6×10^5	ND	ND	ND	3.8	130	3.0×10^4	ND	ND	ND
RAKR	ND	ND	ND	ND	ND	ND	ND[d]	ND[d]	ND[d]	11	23	4.9×10^5
KAKR	ND	ND	ND	ND	ND	ND	ND	ND	ND	9.8	29	3.4×10^5
RGKR	ND	ND	ND	ND	ND	ND	ND	ND	ND	9.8	95	1.0×10^5
AGKR	ND	ND	ND	ND	ND	ND	13	95	1.3×10^5	0.15	510	3.0×10^2
FGKR	ND	ND	ND	ND	ND	ND	4.9	10	4.7×10^5	6.6	46	1.4×10^5
LGKR	ND	ND	ND	ND	ND	ND	15	17	8.7×10^5	2.3	250	8.9×10^3
RGAR	ND	ND	ND	ND	ND	ND	ND	ND	ND	2.3	570	4.1×10^3 [e]
RGKF	ND	ND	ND	ND	ND	ND	ND	ND	ND	0.22	350	6.3×10^2 [f]

[a] Kinetic constants were determined from plots of initial rates versus substrate concentration. Errors in fits were typically less than 10% but as high as 25% in cases when K_m exceeded 500 μM. ND, not determined. [b] Unable to saturate the enzyme; apparent k_{cat}/K_m calculated from rates at low substrate concentrations assuming $v = (k_{cat}/K_m)[E][S]$. [c] Artificially low k_{cat}/K_m due to cleavage between Lys and Arg-pNA. [d] HPLC analysis of products indicated exclusive cleavage between Cys and Arg-pNA. [e] HPLC analysis of products indicated partial (~30%) cleavage between Ala and Arg-pNA. [f] HPLC analysis of products indicated partial (~20%) cleavage between Lys and Phe-pNA.

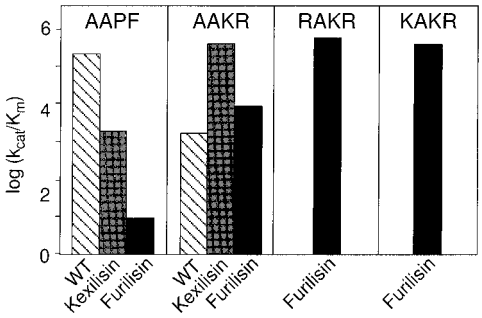

FIGURE 3: Catalytic efficiencies (k_{cat}/K_m) for wild type, kexilisin, and furilisin versus the substrates: succinyl-AAPF-pNA, succinyl-AAKR-pNA, succinyl-RAKR-pNA, and succinyl-KAKR-pNA. Data taken from Table 2.

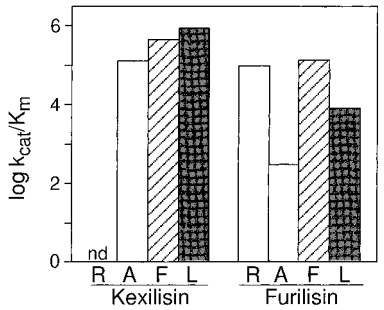

FIGURE 4: Catalytic efficiencies (k_{cat}/K_m) for kexilisin and furilisin versus substrates having variable P4 residues (succinyl-XGKR-pNA). The value for kexilisin versus succinyl-RGKR-pNA was indeterminable due to cleavage between Lys and Arg-pNA. Data taken from Table 2.

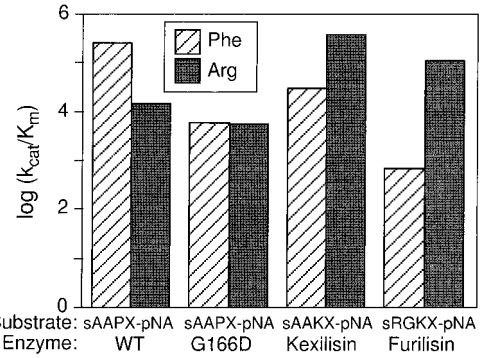

FIGURE 5: Dependence of the P1 specificity for wild-type and variant subtilisins using substrates where the charge at the P2 and P4 positions is varied. Data taken from Table 2.

inhibited with a K_i* of 1.7 nM by the R21M OMTKY3 but not at all by the A15R/T17K/L18R OMTKY3, even up to inhibitor concentrations of 20 μM (Figure 6, Table 3). In contrast, furilisin is inhibited with a K_i* of 29 nM by the A15R/T17K/L18R variant, and the R21M inhibitor gave no detectable inhibition up to concentrations of 3 μM.

DISCUSSION

Modular Nature of Specificity. Multiple subsite recognition allows proteases to achieve exquisite substrate specificities. For example, trypsin, which has specificity for basic residues at the P1 position, should theoretically cleave an average of once every 7.9 residues (based on natural frequency of occurrence; Klapper, 1977). Furin, on the other hand, having specificity for P1Arg, P2Lys/Arg/Pro, and P4Arg substrates (Seidah et al., 1991; Matthews et al., 1994) would cleave once every 2600 amino acids. In our efforts to design site-specific subtilisin variants, we have likewise attempted to accumulate substrate specificity via engineering at multiple subsites. The basic amino acids represent good targets for specificity design due to their high surface accessibility resulting from their positive charge.

An unexpected property of the modular specificity design was the apparent synergy between the engineered subsites. In particular, the status of the S4−P4 interaction affected the S1−P1 site specificity, much the way the S2−P2 interaction did in the preceding work (Ballinger et al., 1995). The result attests to interactions between the S1 and S4 subsites (Grøn & Breddam, 1992). The cumulative effect of the three furilisin mutations on P1 Arg versus Phe preference was nearly 3000-fold.[3]

Charged mutations can exert significant long-range effects on catalytic groups and specificity. For example, in subtilisin the p*K*a of the active site histidine was influenced by ~0.4 pH unit (i.e., ~0.5 kcal/mol) for surface charge perturbations 12−15 Å away, and these substitutions also had a substantial impact on the P1 charge specificity (Russell & Fersht, 1987). For comparison, distances between modeled C_γ atoms of Asp62 and Asp104 to Asp166 in furilisin are 16 and 12 Å,

[3] Assuming a negligible effect of the P3 (Ala or Gly) on the P1 specificity, since the succinyl-AAPX-pNA and succinyl-RGKX-pNA substrate series are being compared.

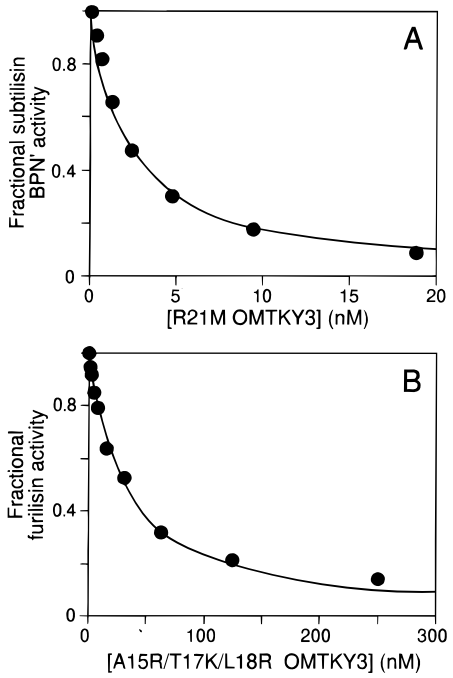

FIGURE 6: Fractional activity for wild-type subtilisin BPN′ versus R21M OMTKY3 (A) and for furilisin versus L15R/T17K/L18R OMTKY3 (B). Incubations were carried out at 21 °C for 20 h in 0.005% Triton X-100 and 20 mM Tris-HCl (pH 8.2) prior to initiating reactions with 75 μM succinyl-AAPF-pNA (wild-type) or 45 μM succinyl-RGKR-pNA (furilisin). Fractional enzyme activities were the average for three determinations at each inhibitor concentration.

Table 3: Inhibition of Wild Type and Furilisin by OMTKY3 Variants[a]

protease	R21M OMTKY3 K_i* (nM)	A15R/T17K/L18R OMTKY3 K_i* (nM)
wild type	1.7 ± 0.1[b]	>20000
furilisin	>3000	29 ± 2
furin[c]	$\sim 5 \times 10^6$ [c]	91[c]

[a] See Materials and Methods for experimental details. Fractional enzyme activities were obtained at each inhibitor concentration in triplicate, and the error shown is the least-squares fit to eq 1 for the average values. [b] The inhibition constant obtained for wild-type subtilisin BPN′ with R21M OMTKY3 is substantially higher than that reported for subtilisin Carlsberg and wild-type OMTKY3 (29 pM; Empie & Laskowski, 1982), which may be due to differences in the specificities of the two subtilisins (Wells et al., 1987). When the buffer conditions were changed to the same as that used by Empie and Laskowski [20 mM $CaCl_2$, 0.005% Triton X-100, 0.2 M Tris-HCl (pH 8.3)], a K_i* of 0.8 nM was measured. [c] Data from Lu et al. (1993). Calculated from the reported K_a values assuming $K_i = 1/K_a$.

respectively. In addition to the long-range electrostatic effects, changes in the S4 site may affect the structure of the S1 site. For example, Rheinnecker et al. (1993, 1994) compared cleavage of tetrameric peptide and acetyltyrosine ethyl ester substrates and found that uncharged mutations in the S4 pocket affected the hydrolysis rates for a substrate that binds only to the S1 subsite. A structure of an inhibitor−enzyme complex should shed light on this issue.

Comparison of Furilisin Mutations to Furin Specificity Determinants. It is interesting to compare the effects of engineering furin-like mutations into subtilisin BPN′ to information from functional studies and structural predictions for furin and other prohormone convertases. Creemers et al. (1993) systematically replaced acidic residues that were predicted to reside near the basic-specific subsites in furin and assayed for cleavage of natural or mutant forms of a prohormone substrate, pro-vWF. Their results implicated Asp165 (S1), Asp33 (S2), Asp104 (S4), and Glu107(S4) (subtilisin BPN′ numbering) as important determinants for the respective subsites, although combined mutations at putative S4 residues E129N, D130N, and D131N also abolished enzyme function. In the S1 subsite, sequence alignments and homology models of human furin (Siezen et al., 1994) and PC2/PC3 (Lipkind et al., 1995) suggest that these enzymes are better optimized for ArgP1 binding than furilisin due to an insertion in the loop at the base of the cavity. Lipkind et al. (1995) indicate that the replacement of a *cis*-proline residue at position 168 in subtilisin BPN′ with a threonine in the prohormone convertases is crucial to allowing Asp165 of the PC's to interact with a P1 Arg residue, since Val165 is oriented away from the subsite cavity in subtilisin BPN′. The G166D mutation in kexilisin and furilisin provides an approximation of this situation although the spacing is not as optimal, since modeling a hydrogen bond to a P1Arg requires that both side chains point out of the cleft (Figure 1B). A basic P1 residue could also interact with Glu156 (Robertus et al., 1972; Poulos et al., 1976). In the S2 pocket, the Asp62 substitution corresponds to the kex2 but not the furin sequence. Furin and the other PC's appear to rely on Asp33 for basic P2 preference, which is not absolute (Molloy et al., 1992). In S4, although the Glu107 of the furin and the PC enzymes can be modeled such that it forms an ion pair with a P4Arg residue (Siezen et al., 1994; Lipkind et al., 1995), we cannot conclude whether this mutation in subtilisin BPN′ could form this interaction because these mutants did not express. Other modifications of the 107 or 126 side chains in furilisin could potentially lead to better discrimination against aromatic P4 residues. The detrimental effects of substitutions at these positions suggest that the subtilisin BPN′ and the prohormone convertase S4-defining segments may differ in backbone structure. Alternatively, optimal cooperation of the 104, 107, and 126 side chains to form a tailor-made Arg-binding pocket may require influence from neighboring portions of the protein.

Comparison to Previous S4 Site Engineering. The Y104D mutation, albeit in the kexilisin background, clearly goes a long way toward basic preference at the S4 subsite. Interestingly, a similar mutation (V104D) in subtilisin 309 did not dramatically alter the ~20-fold Ala over ArgP4 preference of the wild-type enzyme (Bech et al., 1992). The difference between these results and those here may be due to differences between subtilisins 309 and BPN′ as well as the mutated S1 and S2 subsite background.

It has been suggested that the 104 side chain is flexible and should allow more than one binding mode for various P4 residues (McPhalen & James, 1988; Takeuchi et al., 1991). From modeling the furilisin mutations and an Arg-Ala-Lys-Arg peptide onto the wild-type, Ala-Ala-Pro-boroPhe-bound subtilisin structure, it is possible to orient the P4 side chain such that it protrudes directly into the hydrophobic cavity generated by Ile107 and Leu126 (i.e., for hydrophobic substrates) or curls around and hydrogen bonds with Asp104 (in the case of Lys or Arg; Figure 1B). The Y104D mutation also reduced the k_{cat} for hydrolysis of AlaP4 substrates much the same way as other mutations that increased the size of the S4 cavity (Rheinnecker et al., 1993, 1994; Sørensen et

al. 1993). This indicates that tight S4−P4 interactions play a more important role in stabilizing the transition state than the ground state.

Furilisin represents a strong progression toward a multiple-site specific enzyme that could be utilized for site-directed proteolysis. The enzyme and/or subsequent generations may be useful as a tool to selectively cleave at engineered tribasic linker sequences between proteins and fused affinity tags. In addition, isolation of individual domains from proteins that require expression of neighboring modules could be facilitated by furilisin.

ACKNOWLEDGMENT

We thank Dr. Michael Laskowski of Purdue University for supplying OMTKY3 inhibitors, Wayne Anstine for assistance with figures, and the Genentech oligonucleotide synthesis group for synthetic DNA.

REFERENCES

Ballinger, M. D., Tom, J., & Wells, J. A. (1995) *Biochemistry 34*, 13312−13319.
Bech, L. M., Sørensen, S. B., & Breddam, K. (1992) *Eur. J. Biochem. 209*, 869−874.
Bech, L. M., Sørensen, S. B., & Breddam, K. (1993) *Biochemistry 32*, 2845−2852.
Beck, J. T., Marsters, S. A., Harris, R. J., Carter, P., Ashkenazi, A., & Chamow, S. (1994) *Mol. Immun. 31*, 1335−1344.
Bieth, J. (1974) in *Proteinase Inhibitors*, pp 463−469, Springer-Verlag, New York.
Bravo, D. B., Gleason, J. B., Sanchez, R. I., Roth, R. A., & Fuller, R. S. (1994) *J. Biol. Chem. 269*, 25830−25837.
Bresnahan, P. A., Leduc, R., Thomas, L., Thorner, J., Gibson, H. L., Brake, A. J., Barr, P. J. & Thomas, G. (1990) *J. Cell Biol. 111*, 2851.
Bryan, P., Wang, L., Hoskins, J., Ruvinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G., & Gallagher, T. (1995) *Biochemistry 34*, 10310−10318.
Carter, P. (1990) in *Protein Purification: From Molecular Mechanisms to Large-Scale Processes*, pp 181−193, American Chemical Society,Washington, DC.
Carter, P., & Wells, J. A. (1987) *Science 237*, 394−399.
Carter, P., & Wells, J. (1988) *Nature 322*, 546−568.
Creemers, J. W. M., Siezen, R. J., Roebroeck, A. J. M., Ayoubi, T. A. Y., Huylebroeck, D., & Van de Ven, W. J. M. (1993) *J. Biol. Chem. 268*, 21826−21834.
Empie, M. W., & Laskowski, M. (1982) *Biochemistry 21*, 2274−2284.
Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., & Wells, J. A. (1986) *Science 233*, 659−663.
Freifelder, D. (1982) in *Physical Biochemistry*, 2nd ed., p 501, W. H. Freeman and Co., New York.
Gallagher, T., Gilliland, G., Wang, L., & Bryan, P. (1995) *Structure 3*, 907−914.
Grøn, H., & Breddam, K. (1992) *Biochemistry 31*, 8967−8971.
Grøn, H., Meldal, M., & Breddam, K. (1992) *Biochemistry 31*, 6011−6018.
Hosaka, M., Nagahama, M., Kim, W.-S., Watanabe, T., Hatsuzakawa, K., Ikemizu, J., Murakami, K., & Nakayama, K. (1991) *J. Biol. Chem. 266*, 12127−12130.
Klapper, M. H. (1977) *Biochem. Biophys. Res. Commun. 78*, 1018−1024.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331−358.
Kunkel, T. A., Bebenek, K., & McClary, J. (1991) *Methods Enzymol. 204*, 125−139.
LaVallie, E. R., & McCoy, J. M. (1995) *Curr. Opin. Biotechnol. 6*, 501−506.
Lipkind, G., Gong, Q., & Steiner, D. F. (1995) *J. Biol. Chem. 270*, 13277−13284.
Lu, W., Zhang, W., Molloy, S. S., Thomas, G., Ryan, K., Chiang, Y., Anderson, S. & Laskowski, M. (1993) *J. Biol. Chem. 268*, 14583−14585.
Matsubara, H., Kasper, C. B., Brown, D. M., & Smith, E. L. (1965) *J. Biol. Chem. 240*, 1125−1130.
Matthews, D. J., Goodman, L. J., Gorman, C. M., & Wells, J. A. (1994) *Protein Sci. 3*, 1197−1205.
McPhalen, C. A., & James, N. G. (1988) *Biochemistry 27*, 6582−6598.
Mizuno, K., Nakamura, T., Ohshima, T., Tanaka, S., & Matsuo, H. (1988) *Biochem. Biophys. Res. Commun. 156*, 246−254.
Molloy, S. S., Bresnahan, P. A., Leppia, S. I., Klimpel, K. R., & Thomas, G. (1992) *J. Biol. Chem. 267*, 16396−16402.
Morrison, J. F. (1969) *Biochim. Biophys. Acta 185*, 269−286.
Nilsson, B., Forsberg, G., Moks, T., Hartmanis, M., & Uhlén, M. (1992) *Curr. Opin. Struct. Biol. 2*, 569.
Perona, J. J., & Craik, C. S. (1995) *Protein Sci. 4*, 337−360.
Poulos, T. L., Alden, R. A., Freer, S. T., Birktoft, J. J., & Kraut, J. (1976) *J. Biol. Chem. 251*, 1097−1103.
Power, S. D., Adams, R. M., & Wells, J. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3096−3100.
Rheinnecker, M., Baker, G., Eder, J., & Fersht, A. R. (1993) *Biochemistry 32*, 1199−1203.
Rheinnecker, M., Eder, J., Pandey, P. S., & Fersht, A. R. (1994) *Biochemistry 33*, 221−225.
Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C., & Wilcox, P. E. (1972) *Biochemistry 11*, 2439−2449.
Russell, A. J., & Fersht, A. R. (1987) *Nature 328*, 496−500.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
Seidah, N. G., Day, R., Marcinkiewicz, M., Benjannet, S., & Chretien, M. (1991) *Enzyme 45*, 271−284.
Shinde, U., & Inoue, M. (1994) *J. Biochem. 115*, 629−636.
Shinde, U., & Inoue, M. (1995) *J. Mol. Biol. 247*, 390−395.
Siezen, R. J., de Vos, W. M., Leunissen, A. M., & Dijkstra, B. W. (1991) *Protein Eng. 4*, 719−737.
Siezen, R. J., Creemers, J. W. M., & Van de Ven, W. J. M. (1994) *Eur. J. Biochem. 222*, 255−266.
Siezen, R. J., Leunissen, J. A. M., & Shinde, U. (1995) in *Intramolecular Chaperones and Protein Folding*, pp 233−256, Springer, New York.
Smeekens, S. P. (1993) *Bio/Technology 11*, 182−186.
Sørensen, S. B., Bech, L. M., Meldal, M., & Breddam, K. (1993) *Biochemistry 32*, 8994−8999.
Steiner, D. F., Smeekens, S. P., Ohagi, S., & Chan, S. J. (1992) *J. Biol. Chem. 267*, 23435−23438.
Takeuchi, Y., Satow, Y., Nakamura, K. T., & Mitsui, Y. (1991) *J. Mol. Biol. 221*, 309−325.
Teplyakov, A. V., van der Laan, J. M., Lammers, A. A., Kelders, H., Kalk, K. H., Misset, O., Mulleners, L. J. S. M., & Dijkstra, B. W. (1992) *Protein Eng. 5*, 413−420.
Uhlen, M., & Moks, T. (1990) *Methods Enzymol. 185*, 129−143.
Wells, J. A., Ferrari, E., Henner, D. J., & Chen, E. Y. (1983) *Nucleic Acids Res. 11*, 7911−7925.
Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estell, D. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5167−5171.
Williams, J. W., & Morrison, J. F. (1979) *Methods Enzymol. 63*, 437−467.
Wise, R. J., Baar, P. J., Wong, P. A., Kiefer, M. C., Brake, A. J., & Kaufman, R. J. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9378−9382.
Yang, M. Y., Ferrari, E., & Henner, D. J. (1984) *J. Bacteriol. 160*, 15−21.

BI961543H