Engineering Subtilisin into a Fluoride-Triggered Processing Protease Useful for One-Step Protein Purification†

Biao Ruan, Kathryn E. Fisher, Patrick A. Alexander, Viktoriya Doroshko, and Philip N. Bryan*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850

Received August 23, 2004; Revised Manuscript Received October 13, 2004

ABSTRACT: Subtilisin was engineered into a highly specific, processing protease, and the subtilisin prodomain was coengineered into an optimized recognition sequence. This involved five steps. First, a robust subtilisin mutant was created, which could tolerate the subsequent mutations needed for high specificity. Second, the substrate binding pocket was mutated to increase its sequence selectivity. Third, the subtilisin prodomain was engineered to direct cleavage to the junction of any protein fused to it. Fourth, the active site of subtilisin was engineered to kinetically isolate binding and cleavage reactions. Finally, specific anions were identified to trigger the processing reaction, with fluoride ions being particularly useful. The ability to isolate the binding and processing steps with a triggering mechanism created a protease with a virtual on−off switch. This allowed column-immobilized processing subtilisin to be used as both the affinity ligand and processing protease for one-step purification of proteins. Fusion proteins tagged with the engineered prodomain can be bound to the column and washed free of contaminants. Cleavage can be triggered by the addition of fluoride to release the pure target protein. The column is then regenerated by stripping off the tightly bound prodomain at pH 2.1. Ten proteins have been purified to date by this method.

Proteases range in function from broad-specificity, degradative enzymes to high-specificity, processing enzymes which regulate physiological processes from embryonic development to cell death (*1*). Processing proteases are also valuable research tools which aid in the purification and analysis of proteins in a manner somewhat analogous to the use of restriction endonucleases in the analysis of DNA. Proteomic and structural genomic initiatives have created a great need for better protein tools (*2*), but natural processing proteases generally are not as robust and versatile as restriction endonucleases. Thus, the ability to engineer a robust bacterial protease, such as subtilisin, to cleave defined sequences is important both technologically and for understanding how proteases regulate physiological processes (*3*).

Subtilisin BPN′ from *Bacillus amyloliquefaciens* is a secreted, serine protease which degrades proteins in the extracellular environment and provides amino acids to the *Bacillus* (*4*). Consequently, the natural enzyme is fairly indiscriminate in its substrate preferences. Subtilisin's broad preferences result in part from the manner in which it binds to protein substrates. Most subtilisin contacts are with the first four substrate amino acids on the acyl side of the scissile bond. These residues are denoted P1−P4, numbering from the scissile bond toward the N-terminus of the substrate (*5*−*7*). The backbone of the substrate inserts between strands 100−104 and 125−129 of subtilisin to become the central strand in an antiparallel β-sheet arrangement involving seven main chain H-bonds (*6*, *8*). Hence, a major component of substrate binding energy involves only the peptide backbone and not side chain interactions. This contrasts with the prohormone-processing subtilisins, furin and kex2 which lack the structural equivalent of the 100−104 strand (*9*, *10*). The side chain components of substrate binding to subtilisin result primarily from the P1 and P4 amino acids (*11*−*13*). Optimal substrates for subtilisin have large hydrophobic amino acids at these positions (*11*, *12*). Subtilisin exhibits little discrimination for sequences on the leaving group side of the scissile bond.

The specificity of a protease can be assessed by its catalytic efficiency (k_{cat}/K_m) for a cognate sequence relative to nonspecific sequences. The catalytic efficiency of a processing protease for its cognate can be as high as 1×10^8 M^{-1} s^{-1} (*14*), approaching the rate of substrate binding. Subtilisin's catalytic efficiency exceeds 10^7 M^{-1} s^{-1} for a preferred sequence but remains $\geq 10^6$ M^{-1} s^{-1} even if it deviates considerably from the optimum, resulting in its broad specificity (*12*). Over the past 20 years, considerable effort has been spent in engineering subtilisin to increase sequence selectivity (*15*). Most approaches have concentrated on substrate binding through reconfiguration of the P1 and P4 binding pockets. The basic dilemma for the protein engineer is how to uncode activity for nonspecific sequences without attenuating catalytic efficiency against a preferred sequence. As described above, the mode of substrate recognition by subtilisin makes engineering high specificity based on selective substrate binding alone difficult. In addition, rapid acylation results in strong coupling between binding and subsequent chemical steps and attenuates the effect of

† This work was supported by NIH Grant GM42560.

* To whom correspondence should be addressed. Telephone: (301) 738-6220. Fax: (301) 738-6255. E-mail: bryan@umbi.umd.edu.

10.1021/bi048177j CCC: $27.50
Published on Web 10/28/2004

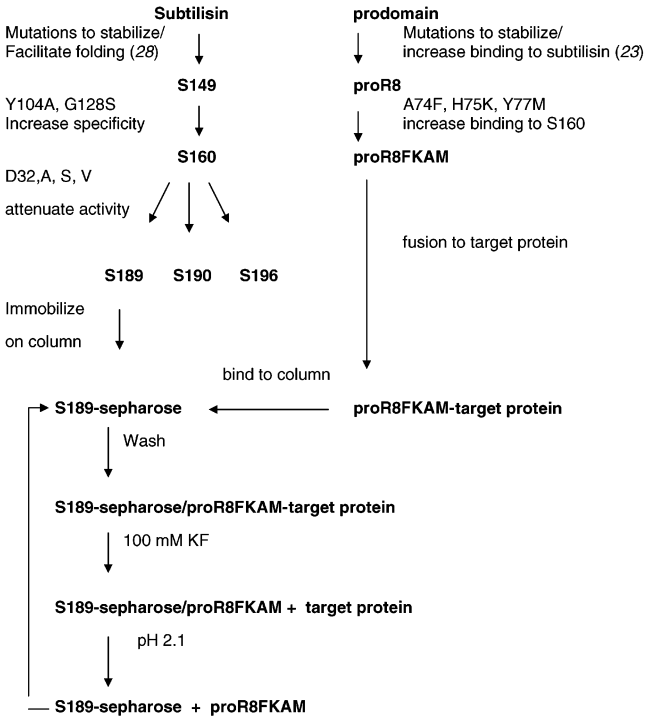

FIGURE 1: Flowchart for the engineering of subtilisin into a fluoride-triggered processing protease and the co-engineering of its prodomain into an optimized cognate sequence. The components are combined to create a one-step purification system for proteins tagged with the cognate prodomain.

differential substrate binding on the overall rate of catalysis. Natural processing proteases achieve specificity by the influence of preferred substrates on both k_{cat} and K_m (*1*). In some highly selective enzymes, binding of a cognate sequence can induce conformational changes in the active site which trigger catalysis (*16*). Carter and Wells previously were able to engineer a profound substrate influence on k_{cat} by mutating the catalytic His 64 to Ala in subtilisin and using a histidine at the P2 position of the substrate as a general base in catalysis (*3*, *17*, *18*). The enzyme is virtually inert against substrates without a P2 histidine. Although only weakly active against even cognate sequences ($k_{cat}/K_m \sim 1000\ M^{-1}\ s^{-1}$), a H64A subtilisin nevertheless proved to be sufficiently useful to become a commercial product (*17*).

Natural processing proteases achieve sequence specificity in subtle ways which have proven problematic to engineer into subtilisin. Here we describe the conversion of subtilisin to a high-specificity, processing enzyme by engineering the enzyme and by recruiting the subtilisin prodomain to help direct specific cleavage. This required five steps. First, a highly robust mutant (S149) was created, which allows subsequent specificity mutations to be tolerated. Second, the P4 binding pocket of S149 was mutated to increase its sequence selectivity. Third, the subtilisin prodomain was engineered into an optimized recognition sequence which directs cleavage to the junction of any protein fused to it. Fourth, the catalytic activity of S149 was attenuated to kinetically isolate binding and cleavage reactions. Finally, specific anions were identified to trigger the cleavage reaction. This specialized processing system is further shown to allow one-step affinity purification of proteins fused to the engineered prodomain. A flowchart of the experimental process is shown in Figure 1.

Table 1: Mutants of Subtilisin

mutation[a]	S149	S160	S189	S190	S196
A(pro74)F			**X**	**X**	**X**
Y(pro77)M	X	X	X	X	X
Q2K	X	X	X	X	X
S3C	X	X	X	X	X
P5S	X	X	X	X	X
S9A	X	X	X	X	X
I31L	X	X	X	X	X
D32A			**X**		
D32S				**X**	
D32V					**X**
K43N	X	X	X	X	X
M50F	X	X	X	X	X
A73L	X	X	X	X	X
Δ75−83	X	X	X	X	X
Y104A		**X**	**X**	**X**	**X**
G128S		**X**	**X**	**X**	**X**
E156S	X	X	X	X	X
G166S	X	X	X	X	X
G169A	X	X	X	X	X
S188P	X	X	X	X	X
Q206C	X	X	X	X	X
N212G	X	X	X	X	X
Y217L	X	X	X	X	X
N218S	X	X	X	X	X
T254A	X	X	X	X	X
Q271E	X	X	X	X	X

[a] Mutations in bold are specific to processing subtilisins.

METHODS

Expression and Purification of Subtilisins. The subtilisin gene from *B. amyloliquefaciens* (subtilisin BPN′) has been cloned, sequenced, and expressed at high levels from its natural promoter sequences in *Bacillus subtilis* (*4*, *19*). All mutant genes (encoding the signal peptide, prodomain, and mature subtilisin) were recloned into a pUB110-based expression plasmid and used to transform *B. subtilis*. Genes for mutant subtilisins also included mutations in the prodomain to increase production of mature subtilisin, as specified in Table 1. Mutant subtilisins S189, S190, and S194 could not be expressed unless the A(pro74)F[1] mutation in the prodomain was present. The *B. subtilis* strain used as the host contains a chromosomal deletion of its subtilisin gene and therefore produces no background wild-type activity (*20*). Subtilisin variants were expressed in a 5 L New Brunswick fermentor at a level of ≥200 mg/L. When the A(pro74)F, D32A, Y104A, and G128S mutations (Table 1) were introduced into subtilisin BPN′, the mutant protein was not secreted, even in the presence of active "helper" subtilisin (*21*). Fermentation conditions and purification were as described previously (*22*).

Site-Directed Mutagenesis. Site-directed mutations were made using QuikChange (Stratagene) mutagenesis kits according to the manufacturer's instructions. S149 is subtilisin BPN′ with Q2K, S3C, P5S, S9A, I31L, K43N, M50F, A73L, Δ75−83, E156S, G166S, G169A, S188P, Q206C, N212G, Y217L, N218S, T254A, and Q271E mutations. S160 is S149 with Y104A and G128S mutations. S189, S190, and S196

[1] A shorthand for denoting amino acid substitutions employs the single-letter amino acid code as follows. A(pro74)F denotes the change of alanine to phenylalanine at postion 74 in the prodomain of subtilisin. Y104A denotes the change of tyrosine to alanine at position 104 of mature subtilisin.

are S160 with D32A, D32S, and D32V mutations, respectively (Table 1). The expression plasmid for mutant prodomains has previously been described (*23*). ProR8 is the 77-amino acid subtilisin BPN prodomain with the following mutations: amino acids 16−21 (QTMSTM) replaced with a SGIK sequence, A23C, K27Q, V37L, Q40C, H72K, and H75K. ProR8FKAM is proR8 with amino acids 74−77 (AKAY) replaced with a FKAM sequence.

Kinetic Measurements. Steady state assays of the hydrolysis of sAAPF−pNA[2] were as described by DelMar et al. (*24*). sDFKAM−AMC was purchased from AnaSpec Inc. The concentration of the AMC substrate was determined by the absorbency at 324 nm using an extinction coefficient of 16 mM^{-1} cm^{-1}. Reaction kinetics of AMC substrates were measured using a KinTek stopped-flow model SF2001 instrument (excitation $\lambda = 380$ nm, emission with a 400 nm cutoff filter). A stock solution of the enzyme at a concentration of 2 mM in 10 mM KPO_4 (pH 5.7), 100 mM NaCl, and 50% glycerol was prepared for kinetic studies. Immediately prior to stopped-flow mixing, the enzyme was diluted into 100 mM KPO_4 (pH 7.2) and placed in syringe A of the SF2000 instrument. A solution of the substrate in the same buffer was place in syringe B. Solutions were equilibrated at 25 °C. Fluorescence data were collected after 1:1 mixing of the two solutions.

Processing Kinetics. Kinetics in 0.1 M KPO_4 at pH 7.2 and 25 °C were determined by binding 2 mg of ProR8-FKAM−G_B to a 1 mL S189HiTrap NHS column and washing into the phosphate buffer. At time intervals, 2 mL of buffer was washed through the column and collected, and the G_B concentration in the aliquot was determined by the absorbance at 278 nm. The cumulative fraction of G_B cleaved was plotted as a function of time to determine a first-order rate constant for hydrolysis.

Expression of Fusion Proteins. A fusion protein vector, pG58, was constructed so that genes of target proteins could be cloned and expressed. The plasmids used as the starting points, pG5 and proR8, have been described previously (*23*, *25*). Unique SnaBI, BbsI, and DraI sites were created near the 3′ end of the proR8FKAM gene to use as 5′ acceptor sites for the fusion of target genes. The multiple cloning block of pG5 is used as 3′ acceptor sites for target genes. Genes were cloned into the pR58 vector and expressed in BL21 DE3 cells essentially as described previously (*25*). Growth and induction conditions varied with the solubility of the specific fusion protein. Cells were harvested by centrifugation at 10000*g* for 15 min, resuspended in 0.1× PBS (pH 7.2), 30 μg/mL DNase I, 5 mM $MgCl_2$, and 0.1 mM PMSF, and lysed by sonication on ice. Cellular debris was pelleted by centrifugation at 10000*g* for 15 min, and the supernatant was delipidated by centrifugation at 100000*g* for 1 h. The high-speed centrifugation step is not mandatory but extends the life of columns.

Purification of the Target Protein. One milliliter of a 20 mg/mL solution of S189 in 0.2 M $NaHCO_3$ (pH 8.3) and 0.5 M NaCl was injected onto a 1 mL *N*-hydroxysuccinamide−Sepharose HiTrap column (Amersham Biosciences) and incubated for 16 h at 4 °C. Washing and deactivation of excess active groups were carried out according to the manufacturer's instructions. To purify GFP, 1 mL of soluble cell extract from 10 mL of a proR8FKAM−GFP culture was injected on the S189 column at a rate of 1 mL/min for 3 min to allow binding and then washed at a rate of 5 mL/min for 2 min to remove impurities. To cleave and elute purified GFP, 2 mL of 100 mM KF was injected at a rate of 0.1 mL/min. To strip and regenerate the column, 1.5 mL of 0.1 N H_3PO_4 was injected at a rate of 1 mL/min. Columns can be recycled at least 30 times.

RESULTS AND DISCUSSION

Engineering a Highly Stable Subtilisin. The starting point for engineering a processing subtilisin was a mutant denoted S149 (Table 1). S149 lacks the calcium-binding loop of natural subtilisins and is highly stable independent of calcium binding. Its engineering has been described in detail in previous publications from our laboratory (*26*−*28*). The robust framework of S149 was essential because the collective mutations to the substrate binding pockets and active site required for high specificity apparently resulted in folding defects when they were introduced into wild-type subtilisin (see Methods). Also, S149 and its variants are able to refold in vitro with high efficiency. As will be discussed below, this property is highly desirable in using immobilized processing subtilisin for protein purification applications.

Increasing Sequence Selectivity. Optimal substrates for subtilisin have F, L, I, or V at P4 (*13*, *29*), A or S at P2 (*13*, *30*) (B. Ruan and P. N. Bryan, manuscript in preparation), and Y, F, L, or M at P1 (*11*, *13*). Rheinnecker et al. (*29*) have previously noted that expanding the P4 pocket can result in a significant increase in specificity by affecting both substrate binding and k_{cat}. They suggested that the expanded P4 cavity causes structural changes which are propagated to the activity site. Optimal active site geometry is restored by filling the cavity with a bulky hydrophobe at the P4 position of a substrate. Consequently, the G128S and Y104A mutations were introduced into S149 (denoted S160, Table 1) to enlarge the P4 pocket (Figure 2A) (*12*, *29*). Steady state kinetic behavior of S149 and S160 was consistent with the earlier observations. The k_{cat}/K_m of S149 versus the peptide substrate sAAPF−pNA is 1.3×10^6 M^{-1} s^{-1} (0.1 M KPO_4 at pH 7.2 and 25 °C). The k_{cat}/K_m of S160 is 80 times smaller (1.6×10^4 M^{-1} s^{-1}). Against an optimized peptide substrate (sDFKAM−AMC), the activity of the two enzymes is similar and is largely determined by the rate of substrate binding (k_{cat}/K_m of S149 = 6.3×10^6 M^{-1} s^{-1} and k_{cat}/K_m of S160 = 4.5×10^6 M^{-1} s^{-1}). A detailed analysis of S160 using transient state kinetic methods will be reported elsewhere (B. Ruan, K. E. Fisher, P. A. Alexander, and P. N. Bryan, manuscript in preparation). The preferred P4−P1 sequence of FKAM is termed a proto-cognate sequence for S160 subtilisin. The preference of S160 for the proto-cognate sequence relative to noncognates is improved compared to that of S149, but is not sufficient for high-specificity processing. We therefore decided to radically increase the size of the recognition sequence.

Incorporating the Proto-Cognate Sequence into a Stabilized Prodomain. The next step was to create a fully cognate recognition sequence by recruiting the subtilisin prodomain to help direct cleavage. The 77-amino acid prodomain is a

[2] Abbreviations: sAAPF−pNA, succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide; sDFKAM−AMC, succinyl-L-Asp-L-Phe-L-Lys-L-Ala-L-Met-7-amino-4-methylcoumarin.

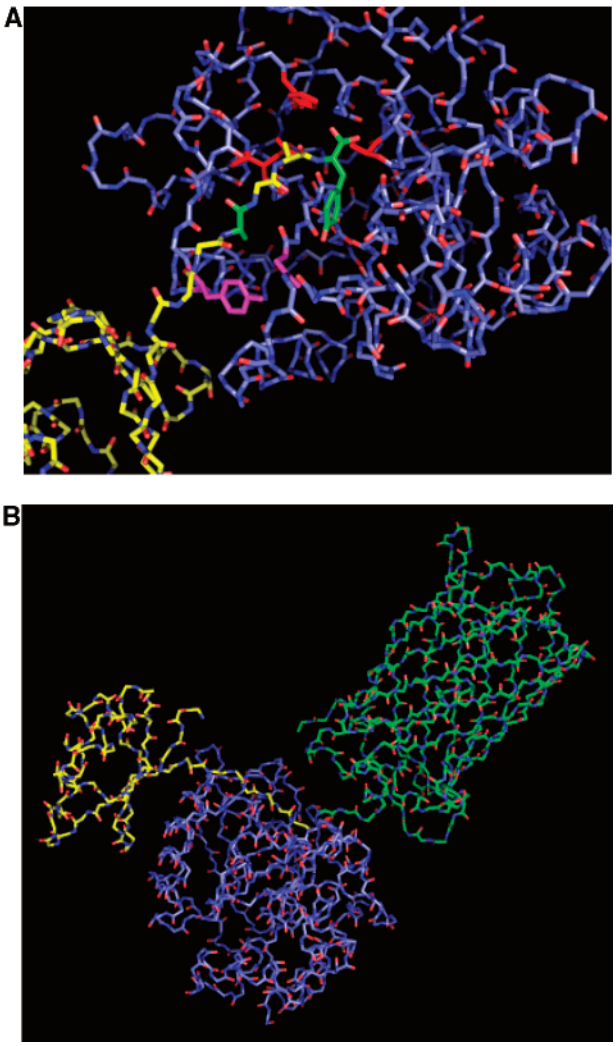

FIGURE 2: Complexes of subtilisin with the prodomain substrate. (A) Backbone depiction of subtilisin (blue) in complex with its prodomain (yellow) [Protein Data Bank entry 1spb (*33*)]. The active site triad is colored red. Prodomain amino acids 72−77 extend into the subtilisin active site to become the central strand in a three-stranded β−sheet. The P1 amino acid (Y77) and P4 amino acid (A74) are colored green. Sites of mutation in the P4 binding pocket of subtilisin (Y104 and G128) are colored violet. (B) Hypothetical model of the proR8FKAM−GFP fusion protein bound to S189 subtilisin. The model is based on the structure of the subtilisin−prodomain complex [PDB entry 1spb (*33*)] and GFP [PDB entry 1gfl (*43*)]. Backbone depictions are of subtilisin (blue), the prodomain (yellow), and GFP (green).

natural, tight-binding recognition sequence which directs sequence-specific autoprocessing during subtilisin biosynthesis (*31*). After autoprocessing, the prodomain remains bound to active subtilisin with nanomolar affinity but is rapidly degraded by free subtilisin to yield a fully active enzyme population. We hypothesized that the prodomain might be used to direct cleavage to the junction of any protein fused to it (Figure 2B). The prodomain in complex with subtilisin is a four-stranded antiparallel β-sheet and two three-turn α-helices (*32*, *33*). The C-terminal prodomain residues extend out from its globular part and bind in a substrate-like manner along subtilisin's active site cleft. Residues 77−74 of the prodomain become P1−P4 substrate amino acids, respectively. The folded prodomain has shape complementarity and high affinity for native subtilisin mediated by both the substrate interactions of the C-terminal tail and a hydrophobic interface provided by the β-sheet (*33*). Essentially, the β-sheet interface of the prodomain is an exo-recognition site which amplifies the P1−P4 binding signal. Optimal binding to S160 required engineering of both the globular and the P1−P4 portions. The globular structure of the prodomain is very unstable when it is separated from subtilisin. Prodomain instability may be a biological requirement which ensures that subtilisin inhibition is transient during biosynthesis but is undesirable for our purposes. We previously devised a phage display procedure for selecting independently stable prodomain mutants (*23*). The selection for stability was based on the thermodynamic linkage between prodomain folding and binding. That is, the correct presentation of the interface with subtilisin depends on the tertiary structure of the prodomain. If mutations are introduced into regions of the prodomain which do not directly contact subtilisin, their effects on binding to subtilisin will be linked to whether they promote the binding conformation. Therefore, mutations which stabilize independent folding of the prodomain increase its binding affinity for subtilisin (*34*). Phagemid selection identified a prodomain, denoted proR8, which is independently stable and binds to subtilisin with ~100 times higher affinity than the wild-type prodomain (*23*). To accommodate the P1−P4 preferences of S160, a mutant of proR8 was constructed with amino acids 74−77 replaced with the proto-cognate FKAM sequence (ProR8FKAM). ProR8FKAM strongly inhibits S160 (K_i = 10 pM). In comparison, the proto-cognate peptide sDFKAM inhibits S160 with a K_i of 360 nM. ProR8FKAM thus became the optimized cognate.

To demonstrate the feasibility of prodomain-directed processing, a gene was constructed to direct the synthesis of a fusion of proR8FKAM directly onto the N-terminus of the 56-amino acid B domain (G_B) of streptococcal Protein G (*25*). The fusion protein (1 μM) was mixed with 1 μM S160. The fusion protein was rapidly and specifically cleaved to release G_B from proR8FKAM. The reaction is effectively terminated after a single catalytic cycle due to product inhibition by proR8FKAM (K_i = 10 pM). The rate of a single cycle of cleavage is limited by the substrate binding rate (1×10^6 M^{-1} s^{-1}). Product inhibition is difficult to avoid in using a high substrate affinity to direct specific cleavage because of the structural similarity between the substrate and product. We did not attempt to obviate this property because the single-turnover reaction can be exploited in using immobilized subtilisin for protein purification. Processing proR8FKAM−G_B with S160 subtilisin is highly specific but only because G_B and the globular part of proR8FKAM are resistant to subtilisin activity. Converting S160 into a useful processing enzyme required a further reduction of its activity against noncognate sequences.

Kinetically Isolating Binding and Cleavage Reactions. Preferential binding of S160 to proR8FKAM fusions relative to noncognate sequences does not, by itself, result in highly specific cleavage because of the kinetic coupling between substrate binding and acylation. This can be discerned by considering the following mechanism for a single catalytic cycle:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} EA + P$$

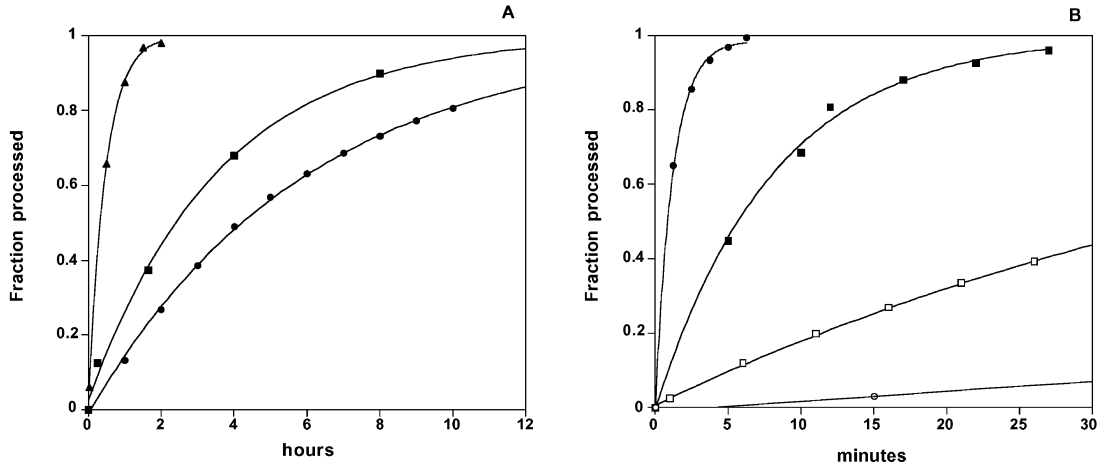

FIGURE 3: Kinetics of cleavage of proR8FKAM−G_B at 25 °C. (A)S160 subtilisin with mutations at D32: A32 (S189) (●), V32 (S196) (■), and S32 (S190) (▲). (B) S189 (A32) as a function of fluoride concentration: 0 (○), 1 (□), 10 (■), and 100 mM (●).

where E is the enzyme, S the substrate, ES the enzyme−substrate complex, EA the acyl enzyme, and P the product released on the leaving group side of the scissile bond. The rate of the release of the product $dP/dt = k_2k_1[S]/(k_1[S] + k_{-1} + k_2)$ (*35*). In the reaction of sDFKAM−AMC with S160, the substrate off rate (k_{-1}) is ∼10 s^{-1} compared to an acylation rate (k_2) of ∼100 s^{-1}. Thus, k_{-1} has little influence on the observed rate of product formation. The same is true for the cognate substrate (proR8FKAM−G_B) ($k_{-1} < 0.0001$ s^{-1}). Substrate affinity would strongly influence cleavage rate, however, if the acylation rate were slow enough that equilibrium between the enzyme and substrate were approximated throughout the reaction cycle. The fourth step in the process, therefore, was to uncouple the binding and cleavage reactions by slowing acylation. This can be accomplished with mutations of catalytic amino acids (D32, H64, or S221) (*36*) or the oxyanion hole amino acid (N155) (*37*, *38*). We focused on mutations at D32 because the rate of acylation (k_2) is slow enough (approximately per hour in 0.1 M KPO_4 at pH 7.2) to create high selectivity for the cognate on the basis of differential binding. It also provides the potential to trigger catalysis by the addition of specific anions, as will be described below. Figure 3A shows the kinetics of cleavage of proR8FKAM−G_B by S160 mutants D32A (S189), D32V (S196), and D32S (S190) at pH 7.2 and 25 °C. As expected, the reactions are slow single-turnover reactions with an 8-fold difference in rate between the slowest (A32, 0.17 h^{-1}) and the fastest (S32, 1.3 h^{-1}). At 4 °C, the cleavage reactions are slowed by ∼4-fold. The high selectivity of the D32 mutants for the cognate proR8FKAM sequence is manifested in several ways. First, they turnover the proto-cognate sDFKAM−AMC peptide very poorly ($k_{cat}/K_m \sim 1$ M^{-1} s^{-1} in 0.1 M KPO_4 at pH 7.2 and 25 °C). Second, they are unable to process the proR8−G_B fusion protein, which lacks the FKAM proto-cognate sequence. Third, they are unreactive with the serine protease inhibitor, phenylmethanesulfonyl fluoride (PMSF). Finally, they are unable to carry out autoprocessing in vivo unless the P4 residue of their prosequence is mutated from A to F. With the A(pro74)F mutation (Table 1), they efficiently autoprocess and are secreted from *Bacillus* at ≥200 mg/L.

Identification of Triggering Anions. The final step was to identify a method to trigger the cleavage reaction. This approach was based on the observation that mutation of an active site residue in an enzyme can sometimes be compensated by addition of a small molecule which mimics the lost functional group (*39*). This was the case for D32 mutants of S160. The carboxylate of D32 hydrogen bonds to the catalytic H64 and allows it to act as a general base first and then a general acid during acylation. In 1987, Craik et al. (*40*) determined that mutation of the catalytic D102 in trypsin drastically decreased peptidase activity below pH 9 but creates an alternative, hydroxide-dependent mechanism which became evident above pH 10. A less pronounced hydroxide-dependent effect was also observed for the D32A mutant of wild-type subtilisin (*21*). We observed that the rate of processing of proR8FKAM−G_B by the D32 mutants was sensitive not only to high pH but also to chloride and fluoride concentrations. Anions such as phosphate, acetate, and carbonate did not accelerate cleavage. Table 2 summarizes the cleavage rates of the D32A mutant (S189) as a function of specific anions at 25 °C. Hydroxide is the most potent triggering anion, but its effects on the ionization state of charged residues, at the concentrations required for rapid triggering, compromise its utility. Fluoride is attractive as a chemical trigger because of its potency and its selective effect on proR8FKAM fusions (Figure 3B). Cleavage of proR8-FKAM−G_B by S189 was accelerated by 280-fold in 100 mM KF, but fluoride affected the turnover of sDFKAM−AMC much less (10-fold increase in k_{cat}/K_m at 100 mM KF). The mechanism of anion activation of S189 and its differential

Table 2: Kinetic Properties of S189 Subtilisin versus proR8FKAM−G_B[a]

pH	[KF] (mM)	[KCl] (mM)	rate (min^{-1})
5.7	0	0	0.002
7.2	0	0	0.003
7.2	1	0	0.018
7.2	10	0	0.14
7.2	100	0	0.8
7.2	0	100	0.01
7.2	0	500	0.05
8.8	0	0	0.016
10	0	0	0.067

[a] First-order rate constants for the release of processed G_B as a function of anion. Reactions in KF and KCl with 0.1 M KPO_4 at pH 7.2 and 25 °C. Reactions vs pH in 0.1 M KPO_4 buffer at 25 °C.

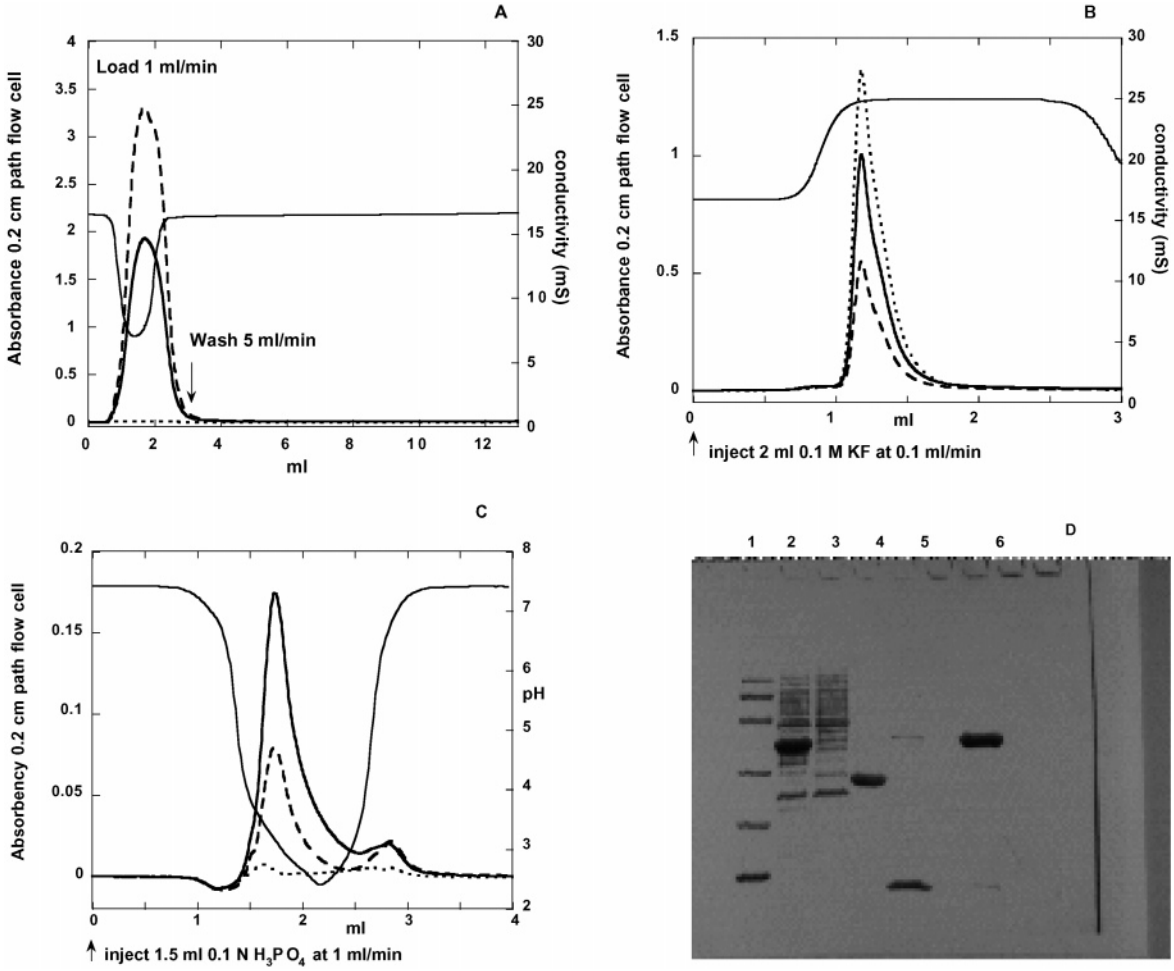

FIGURE 4: Purification of GFP from the fusion protein (proR8FKAM−GFP) on a 1 mL S189HiTrap NHS column. (A−C) Chromatographic traces from the purification: (thick —) A280, (– – –) A250, (•••) A488, and (thin —) conductivity or pH. (A) Loading/washing sequence. (B) Cleave/elute sequence. (C) Strip sequence. (D) SDS−PAGE: lane 1, molecular weight standards (2 μg band); lane 2, cell lysate (5 μL of 5 mL); lane 3, flow-through of a 1 mL loading (10 μL of 2 mL); lane 4, cleavage/elution in 0.1 M KF (5 μL of a 1 mL fraction); lane 5, strip by 0.1 N H_3PO_4 (10 μL of a 1 mL fraction); and lane 6, strip by 0.1 N H_3PO_4 with no KF triggering (10 μL of a 1 mL fraction).

effects on proR8FKAM fusions as compared to peptide substrates is of great practical value but is not yet understood.

Purification of Fusion Proteins. The binding and catalytic properties of S189 subtilisin can be exploited for affinity purification and processing of proteins fused to proR8FKAM. To demonstrate this point, we immobilized 15 mg of S189 on 1 mL of Sepharose beads. An *Escherichia coli* cell lysate containing a fusion of proR8FKAM and green fluorescent protein (GFP) (Figure 2B) was passed over the S189 column at a flow rate of 1 mL/min. The binding of proR8FKAM−GFP to the column can be monitored by the absorbency at 488 nM. The fusion protein is bound rapidly and quantitatively, while the impurities are washed through the matrix (Figure 4A). Release of purified GFP is effected by the addition of a triggering anion (100 mM KF) (Figure 4B). Cleavage can also be effected by extended incubation without a triggering anion [e.g., 15 h in 0.1 M KPO_4 (pH 7.2)]. Purified and correctly processed GFP (as determined by MALDI MS) was recovered quantitatively in the KF eluate. Up to 5 mg of GFP can be purified in a single shot from a 1 mL S189 column. The column can be reused after the tightly bound proR8FKAM is stripped from S189 at pH 2.1 (Figure 4C). The binding of proR8FKAM to S189 is so tight that acid denaturation is required to disrupt the complex. The facile refolding of S189 allows columns to be recycled at least 30 times without a significant decrease in performance, however. Binding and cleavage reactions of proR8FKAM fusions are insensitive to EDTA, reducing agents, detergents, $\leq$2 M urea, and $\leq$1 mM PMSF, although fluoride ions released from spontaneously hydrolyzed PMSF will increase the cleavage rate. To date, 10 proR8FKAM fusion proteins have been constructed and purified by this method (Table 3). The proR8FKAM tag improved the soluble expression of those proteins which were only partially soluble when expressed without the tag (e.g., transducin, arrestin, GFP, and CDC6). All target proteins were correctly processed and recovered in high yield. In all cases, the desired N-terminus

Table 3: Proteins Purified Using S189 Columns

streptococcal protein G_B domain	56 amino acids
streptococcal protein G_a domain	45 amino acids
protein G_B mutant G311	56 amino acids
staphylococcal protein A_B domain	56 amino acids
protein A_B mutant A219	56 amino acids
E. coli hypothetical Yab	117 amino acids
bovine α-subunit of transducin	350 amino acids
Methanobacterium thermautotrophicus CDC6	379 amino acids
Aequorea victoria green fluorescent protein (GFP)	238 amino acids
bovine arrestin	390 amino acids

of the target protein was generated by the processing reaction. Spacer amino acids between proR8FKAM and the target protein are not required because amino acids on the leaving group side of the scissile bond have a minimal influence on binding of the substrate to S189.

CONCLUSIONS

The engineering of site-selective subtilisins and an optimized cognate sequence resulted from a confluence of proficiencies acquired over many years: (1) engineering of highly stable and refoldable subtilisin (*26*, *41*), (2) modifications of the P4 binding pocket to increase sequence selectivity (*12*, *29*, *42*), (3) selection of independently stable prodomains with a high affinity for subtilisin (*23*, *34*), (4) attenuation of catalytic activity to kinetically isolate binding and cleavage steps (*36*−*38*), and (5) identification of triggering anions to compensate for active site mutations (*39*, *40*). The ability to modulate the rate of the cleavage reaction with a combination of anion concentration and temperature creates a processing system with a virtual on−off switch. Detailed kinetic and structural characterization of the binding and processing reaction should allow new generations of enzymes, cognate sequences, and triggering mechanisms to be developed.

ACKNOWLEDGMENT

We thank Feng Hong Song for synthesizing the oligonucleotides used in site-directed mutagenesis and DNA sequencing and Kevin Ridge, Nadik Aboulaev, Sadhana Pullalarevu, Osnat Herzberg, and Zvi Kelman, who subcloned target genes and produced fusion proteins to enable testing of the purification method.

REFERENCES

1. Hedstrom, L. (2002) Serine protease mechanism and specificity, *Chem. Rev. 102*, 4501−4524.
2. Burley, S. K. (2000) An overview of structural genomics, *Nat. Struct. Biol. 7* (Suppl.), 932−934.
3. Carter, P., and Wells, J. A. (1987) Engineering enzyme specificity by "substrate-assisted catalysis", *Science 237*, 394−399.
4. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., and Chen, E. Y. (1983) Cloning, sequencing and secretion of *Bacillus amyloliquifaciens* subtilisin in *Bacillus subtilis*, *Nucleic Acids Res. 11*, 7911−7925.
5. Berger, A., and Schechter, I. (1970) Mapping the active site of papain with the aid of peptide substrates and inhibitors, *Philos. Trans. R. Soc. London, Ser. B 257*, 249−264.
6. McPhalen, C. A., Schnebli, H. P., and James, M. N. (1985) Crystal and molecular structure of the inhibitor eglin from leeches in complex with subtilisin Carlsberg, *FEBS Lett. 188*, 55−58.
7. Bode, W., Papamokos, E., Musil, D., Seemueller, U., and Fritz, H. (1986) Refined 1.2 Å crystal structure of the complex formed between subtilisin Carlsberg and the inhibitor eglin c. Molecular structure of eglin and its detailed interaction with subtilisin, *EMBO J. 5*, 813−818.
8. McPhalen, C. A., and James, M. N. G. (1988) Structural comparison of two serine proteinase-protein inhibitor complexes: Eglin-C-Subtilisin Carlsberg and CI-2-Subtilisin novo, *Biochemistry 27*, 6582−6598.
9. Henrich, S., Cameron, A., Bourenkov, G. P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W., and Than, M. E. (2003) The crystal structure of the proprotein processing proteinase furin explains its stringent specificity, *Nat. Struct. Biol. 10*, 520−526.
10. Holyoak, T., Wilson, M. A., Fenn, T. D., Kettner, C. A., Petsko, G. A., Fuller, R. S., and Ringe, D. (2003) 2.4 Å resolution crystal structure of the prototypical hormone-processing protease Kex2 in complex with an Ala-Lys-Arg boronic acid inhibitor, *Biochemistry 42*, 6709−6718.
11. Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., and Wells, J. A. (1986) Probing steric and hydrophobic effects on enzyme−substrate interactions by protein engineering, *Science 233*, 659−663.
12. Gron, H., and Breddam, K. (1992) Interdependency of the binding subsites in subtilisin, *Biochemistry 31*, 8967−8971.
13. Gron, H., Meldal, M., and Breddam, K. (1992) Extensive comparison of the substrate preferences of two subtilisins as determined with peptide substrates which are based on the principle of intramolecular quenching, *Biochemistry 31*, 6011−6018.
14. Rockwell, N. C., Wang, G. T., Krafft, G. A., and Fuller, R. S. (1997) Internally consistent libraries of fluorogenic substrates demonstrate that Kex2 protease specificity is generated by multiple mechanisms, *Biochemistry 36*, 1912−1917.
15. Bryan, P. N. (2000) Protein engineering of subtilisin, *Biochim. Biophys. Acta 1543*, 203−222.
16. Volanakis, J. E., and Narayana, S. V. (1996) Complement factor D, a novel serine protease, *Protein Sci. 5*, 553−564.
17. Carter, P., Abrahmsen, L., and Wells, J. A. (1991) Probing the mechanism and improving the rate of substrate-assisted catalysis in subtilisin BPN′, *Biochemistry 30*, 6141−6148.
18. Carter, P., Nilsson, B., Burnier, J. P., Burdick, D., and Wells, J. A. (1989) Engineering Subtilisin BPN′ for site-specific proteolysis, *Proteins: Struct., Funct., Genet. 6*, 240−248.
19. Vasantha, N., Thompson, L. D., Rhodes, C., Banner, C., Nagle, J., and Filpula, D. (1984) Genes for alkaline and neutral protease from *Bacillus amyloliquifaciens* contain a large open-reading frame between the regions coding for signal sequence and mature protein, *J. Bacteriol. 159*, 811−819.
20. Fahnestock, S. R., and Fisher, K. E. (1987) Protease-deficient *Bacillus subtilis* host strains for production of staphylococcal Protein A, *Appl. Environ. Microbiol. 53*, 379−384.
21. Carter, P., and Wells, J. A. (1988) Dissecting the catalytic triad of a serine protease, *Nature 332*, 564−568.
22. Alexander, P. A., Ruan, B., and Bryan, P. N. (2001) Cation-dependent stability of subtilisin, *Biochemistry 40*, 10634−10639.
23. Ruan, B., Hoskins, J., Wang, L., and Bryan, P. N. (1998) Stabilizing the subtilisin BPN′ pro-domain by phage display selection: how restrictive is the amino acid code for maximum protein stability?, *Protein Sci. 7*, 2345−2353.
24. DelMar, E., Largman, C., Brodrick, J., and Geokas, M. (1979) A sensitive new substrate for chymotrypsin, *Anal. Biochem. 99*, 316−320.
25. Alexander, P., Fahnestock, S., Lee, T., Orban, J., and Bryan, P. (1992) Thermodynamic analysis of the folding of the streptococcal Protein G IgG-binding domains B1 and B2: why small proteins tend to have high denaturation temperatures, *Biochemistry 31*, 3597−3603.
26. Bryan, P., Alexander, P., Strausberg, S., Schwarz, F., Wang, L., Gilliland, G., and Gallagher, D. T. (1992) Energetics of folding subtilisin BPN′, *Biochemistry 31*, 4937−4945.
27. Strausberg, S., Alexander, P., Gallagher, D. T., Gilliland, G., Barnett, B. L., and Bryan, P. (1995) Directed evolution of a subtilisin with calcium-independent stability, *Bio/Technology 13*, 669−673.
28. Alexander, P. A., Ruan, B., Strausberg, S. L., and Bryan, P. N. (2001) Stabilizing mutations and calcium-dependent stability of subtilisin, *Biochemistry 40*, 10640−10644.
29. Rheinnecker, M., Eder, J., Pandey, P. S., and Fersht, A. R. (1994) Variants of subtilisin BPN′ with altered specificity profiles, *Biochemistry 33*, 221−225.
30. Ballinger, M. D., Tom, J., and Wells, J. A. (1995) Designing subtilisin BPN′ to cleave substrates containing dibasic residues, *Biochemistry 34*, 13312−13319.
31. Ikemura, H., Takagi, H., and Inouye, M. (1987) Requirement of pro-sequence for the production of active subtilisin E in *Escherichia coli*, *J. Biol. Chem. 262*, 7859−7864.
32. Bryan, P., Wang, L., Hoskins, J., Ruvinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G., and Gallagher, T. D. (1995) Catalysis of a protein folding reaction: Mechanistic implications of the 2.0 Å structure of the subtilisin-prodomain complex, *Biochemistry 34*, 10310−10318.
33. Gallagher, T. D., Gilliland, G., Wang, L., and Bryan, P. (1995) The prosegment-subtilisin BPN′ complex: crystal structure of a specific foldase, *Structure 3*, 907−914.

34. Ruvinov, S., Wang, L., Ruan, B., Almog, O., Gilliland, G., Eisenstein, E., and Bryan, P. (1997) Engineering the independent folding of the subtilisin BPN′ prodomain: Analysis of two-state folding vs protein stability, *Biochemistry 36*, 10414−10421.
35. Johnson, K. A. (1992) Transient-state kinetic analysis of enzyme reaction pathways, *Enzymes 20*, 1−61.
36. Carter, P., and Wells, J. A. (1990) Functional interaction among catalytic residues in subtilisin BPN′, *Proteins: Struct., Funct., Genet. 7*, 335−342.
37. Bryan, P., Pantoliano, M. W., Quill, S. G., Hsiao, H. Y., and Poulos, T. (1986) Site-directed mutagenesis and the role of the oxyanion hole in subtilisin, *Proc. Natl. Acad. Sci. U.S.A. 83*, 3743−3745.
38. Wells, J. A., Cunningham, B. C., Graycar, T. P., and Estell, D. A. (1986) Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, *Philos. Trans. R. Soc. London 317*, 415−423.
39. Toney, M. D., and Kirsch, J. F. (1989) Direct Bronsted analysis of the restoration of activity to a mutant enzyme by exogenous amines, *Science 243*, 1485−1488.
40. Craik, C. S., Roczniak, S., Largman, C., and Rutter, W. J. (1987) The catalytic role of the active site aspartic acid in serine proteases, *Science 237*, 909−913.
41. Gallagher, T. D., Bryan, P., and Gilliland, G. (1993) Calcium-free subtilisin by design, *Proteins: Struct., Funct., Genet. 16*, 205−213.
42. Bech, L. M., Sorensen, S. B., and Breddam, K. (1992) Mutational replacements in subtilisin 309. Val104 has a modulating effect on the P4 substrate preference, *Eur. J. Biochem. 209*, 869−874.
43. Yang, F., Moss, L. G., and Phillips, G. N., Jr. (1996) The molecular structure of green fluorescent protein, *Nat. Biotechnol. 14*, 1246−1251.

BI048177J