Biochemistry

 Volume 32, Number 5 February 9, 1993

Accelerated Publications

Engineering a Novel Specificity in Subtilisin BPN′ †

Michael Rheinnecker, Graham Baker,‡ Jörg Eder, and Alan R. Fersht*

MRC Unit for Protein Function and Design, Cambridge Centre for Protein Engineering, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.

Received November 5, 1992; Revised Manuscript Received December 4, 1992

ABSTRACT: The specificity of subtilisin BPN′ toward substrates with large hydrophobic P_4 residues has been improved by single amino acid replacements at positions 104 and 107. Mutations were designed to (i) increase the size of the P_4 binding pocket by replacing Ile107, which is at the bottom of the S_4 pocket, by Val, Ala, and Gly and (ii) lose the hydrogen bond between Tyr104 and Ser130 at the entrance of the P_4 binding pocket by changing Tyr104 to Phe and thus reduce interactions between small P_4 side chains and residue 104. All mutant subtilisins, except for I107V, have increased specificity for residues with large side chains at P_4 compared with wild type. Using the conventional definition of specificity as the competition of different substrates for the same enzyme, the I107G mutant subtilisin has one of the largest improvements in substrate specificity reported for subtilisin so far; k_{cat}/K_M is increased >200-fold for Phe compared with Ala as the P_4 residue. Further, the activity of I107G toward its specific substrate is comparable to that of the wild-type enzyme. Surprisingly, much of the increase in specificity on mutation of Ile107 → Gly appears to result from a lesion that is transmitted through the structure and effects catalysis. The value of k_{cat}/K_M for the small substrate acetyltyrosine ethyl ester, which binds to the S_1 pocket, drops by 93% on mutation of Ile107 → Gly. The lesion in subtilisin I107G is complemented, however, on binding of longer substrates that have a large hydrophobic P_4 amino acid side chain that can bind in the S_4 pocket.

The serine protease subtilisin BPN′ from *Bacillus amyloliquefaciens* has proved to be a good model system for protein engineering studies (Wells, 1987; Wells & Estell, 1988). It fulfills all three requirements for detailed analysis since (i) the high-resolution structures of subtilisin BPN′ (Wright et al., 1969; Kraut, 1977; Bott et al., 1988) and its complex with various polypeptide inhibitors, e.g., eglin *c* (Bode et al., 1986), streptomyces subtilisin inhibitor (SSI) (Mitsui et al., 1979), and chymotrypsin inhibitor 2 (CI2)[1] (McPhalen & James, 1988) have been solved, (ii) suitable cloning and expression systems have been developed (Wells et al., 1983; Thomas et al., 1985), and (iii) accurate methods to quantitate the effect of mutations on proteolytic activity are well established (Thomas et al., 1985; Estell et al., 1986; Grøn et al., 1992).

A major goal in protein engineering is to develop enzymes with new activities or altered specificity. Subtilisin BPN′ binds substrates with a rather broad specificity, mainly by using two specificity pockets, S_1 and S_4, with which the corresponding amino acid side chains P_1 and P_4 of the substrate interact. Previous attempts to change the enzyme to a specific protease by altering the electrostatic (Wells et al., 1987) and hydrophobic and steric environment (Estell et al., 1986) have

† M.R. was supported by the MRC and Schering Agrochemicals Ltd. J.E. was supported by the Deutsche Forschungsgemeinschaft.

‡ Present address: Ferring Research Institute, Southampton University Research Centre, Chilworth, Southampton SO1 7NP, U.K.

[1] Abbreviations: ATEE, acetyltyrosine ethyl ester; Cbz, carbobenzoxy; CI2, chymotrypsin inhibitor 2; MES, 2-(*N*-morpholino)ethanesulfonic acid; NaDODSO$_4$–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis. The binding sites on the enzyme are represented as S_4, S_3, S_2, and S_1 (Schechter & Berger, 1967) while P_4 and P_1 represent the substrate amino acid side chains that interact with the enzyme's binding sites S_4 and S_1. Substrates have the form succinyl-X-Ala-Pro-Phe-*p*-nitroanilide (where X represents the P_4 amino acid) and are abbreviated to s-XAPFpNA. Mutants are designated by the single-letter code of the wild-type amino acid, followed by the residue number and the amino acid replacement.

0006-2960/93/0432-1199$04.00/0

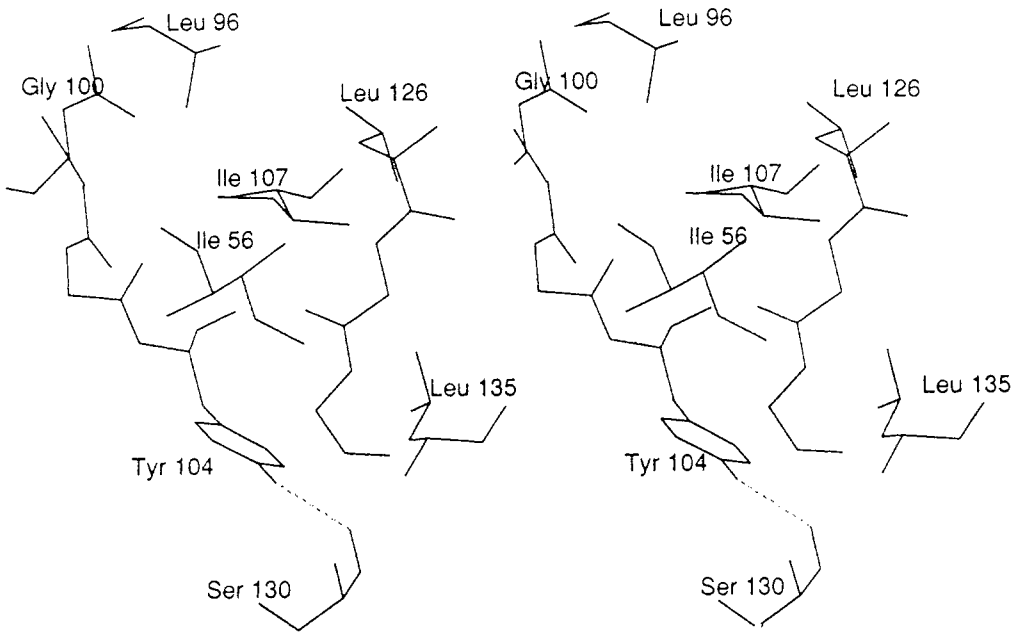

FIGURE 1: Stereographic view of the S_4 specificity pocket of subtilisin BPN′. The view is into the pocket, binding the P_4 residue of chymotrypsin inhibitor 2 (Ile56). Leu96, Gly100, Tyr104, Ile107, Leu126, Ser130, and Leu135 are labeled.

concentrated mainly on the S_1 pocket. Further, using the mechanism of substrate-assisted catalysis, a mutant subtilisin was developed, which was highly specific but showed greatly reduced turnover of substrate (Carter & Wells, 1987; Carter et al., 1991). Recently, Grøn et al. (1992) extended earlier specificity studies by Morihara (1974) and Philipp and Bender (1983) with peptide substrates of varying length and confirmed that the S_4 pocket (Figure 1) is nearly as important as the S_1 pocket in determining the substrate specificity of subtilisin BPN′. In addition, the functional and structural consequences of mutations at the S_4 specificity pocket of the high-alkaline serine protease PB92 from *Bacillus alcalophilus* were recently reported (Teplyakov et al., 1992; van der Laan et al., 1992). We show now that it is possible to alter substantially the substrate specificity of subtilisin BPN′ by mutating residues in the S_4 specificity pocket. In this work, we describe the analysis of a series of mutants (in the S_4 pocket) at positions 107 and 104. The mutant enzymes were designed on the basis of the known crystal structures of subtilisin BPN′ and its complex with CI2 (McPhalen & James, 1988). The largest change in enzyme specificity was obtained by replacing Ile107 with Gly, which resulted in a >200-fold increase in k_{cat}/K_M for Phe as the P_4 side chain compared with Ala in this position.

MATERIALS AND METHODS

Materials. Enzymes for DNA manipulations were from New England Biolabs. Oligonucleotides were synthesized on an Applied Biosystems ABI 380B oligonucleotide synthesizer. The substrates s-AAPFpNA and ATEE were from Sigma. The oligonucleotide-directed in vitro mutagenesis system RPN1523 was obtained from Amersham International plc. CM-trisacryl resin was obtained from IBF Biotechnics. All other reagents were of the highest quality commercially available.

Tetrapeptide Synthesis. The series s-XAPFpNA, where X represents Val, Ile, Leu, and Phe, was synthesized, using a modification of the procedure of Delmar et al. (1979). For the coupling reaction, 3.4 mmol of 1-[3-(dimethylamino)-propyl]-3-ethylcarbodiimide hydrochloride was added to a mixture of 3.1 mmol of CbzAla-Pro, 3.4 mmol of 1-hydroxybenzotriazole hydrate, 3.4 mmol of Phe-*p*-nitroanilide, and 3.4 mmol of triethylamine in 10 mL of dichloromethane. The reaction was stirred under an atmosphere of argon at 0 °C for 1 h and then at room temperature for 24 h. The reaction mixture was washed with equal volumes of 1 M citric acid and saturated sodium hydrogen carbonate solutions. The dichloromethane solution was then dried over anhydrous magnesium sulfate and filtered, and the solvent was removed in vacuo.

For the acid deprotection step, 5 mL of HBr 16% (w/v) in glacial acetic acid was added to the Cbz-peptide product residue, Cbz-APF-*p*-nitroanilide. The reaction was stirred at room temperature for 1 h, and then 50 mL of diethyl ether was added to precipitate the free amino hydrogen bromide salt. The crystals were washed and dissolved in water. The free amino peptide was precipitated by increasing the pH to 10 with 1 M NaOH solution and extracted into dichloromethane. The combined organic fractions were pooled and dried over anhydrous magnesium sulfate. After filtration, the solvent was removed in vacuo. The residue was dissolved in dichloromethane and the product triturated from solution by the addition of diethyl ether to yield white crystals (2.14 mmol, 69% overall).

For the succinylation of the tetrapeptide product, 1.1 mmol of succinyl anhydride in 5 mL of anhydrous dichloromethane was added to 0.7 mmol of NH_2-XAPF-pNA and 2.0 mmol of anhydrous triethylamine in 5 mL of anhydrous dichloromethane under an atmosphere of argon at room temperature. After 24 h, solvent was removed in vacuo and the product was triturated with 50 mL of 0.5 M HCl (aqueous). The crystalline product was filtered and washed with 0.5 M HCl (aqueous) and water. The product was dried overnight under high vacuum and then triturated from THF solution by careful addition of diethyl ether to yield white crystals (65% yield).

For the acetylation, the same methodology as described above was used with acetyl anhydride in place of succinyl anhydride. All reaction products were analyzed with TLC using shadowing at 254 nm. The purity of all peptide substrates was >98% as determined by FAB mass spectrometry and microanalysis.

Mutant Design and Engineering. Using the crystal structure for the CI2–subtilisin BPN′ complex (McPhalen & James,

1988), molecular modeling was performed with the computer program Insight II (Biosym Technologies) on an IRIS workstation. The construction of mutants I107V, I107A, I107G, and Y104F was essentially as described previously (Thomas et al., 1985) except that the mutagenesis reaction was carried out with the Amersham mutagenesis kit. Oligonucleotides used were

Y104F: 5′-GAT CCA GCT GAA TTG GCC GGA-3′

I107V/A/G:
5′-TCC GTT AAT A(A/G/C)C CCA GCT GTA-3′

The I107V/A/G oligonucleotide was a mixture of three nucleotides differing in one base. All mutants were checked by dideoxy sequencing of the entire coding region (Sanger et al., 1977).

Expression and Purification of Wild-Type and Mutant Subtilisins. Wild-type and mutant subtilisins were expressed in the protease-deficient *Bacillus subtilis* strain DB104 (Kawamura & Doi, 1984). The mutant subtilisins were purified from 36-h culture supernatants grown in L-broth supplemented with kanamycin (25 μL/mL), glucose (0.5%), and 2 mM $CaCl_2$. The culture supernatant was concentrated at 4 °C with an Amicon PM10 membrane filter before dialysis against 10 mM MES buffer, pH 6.2, containing 2 mM $CaCl_2$ and 10 mM 2-mercaptoethanol. The protein solution was then loaded onto a column of CM-trisacryl (7.5 cm × 12 cm) and eluted with 10 mM MES buffer, pH 6.2, containing 80 mM NaCl, 2 mM $CaCl_2$, and 10 mM 2-mercaptoethanol. The purified proteins were flash-frozen in bead form and stored at −70 °C. The purity of the proteins was analyzed by $NaDODSO_4$–PAGE.

Kinetic Methods. The kinetic parameters k_{cat} and K_M for the substrates s-XAPFpNA (Delmar et al., 1979; Nakajima et al., 1979) were obtained from initial rate measurements of substrate hydrolysis monitored by following absorption at 412 nm ($\Delta\epsilon_{412}$ = 8480 M^{-1} cm^{-1}). The enzyme assays were performed in 1 mL of 0.1 M Tris-HCl buffer, pH 8.6, at 25 °C with substrate concentrations varying between 0.1 and 2 mM. Reactions were initiated with the addition of 10 μL of the corresponding enzyme solution (4–7 μM).

The k_{cat}/K_M values for the substrate ATEE (Schwert & Takenaka, 1955) were obtained from pseudo-first-order rate measurements monitored by following absorption at 237 nm ($\Delta\epsilon_{237}$ = −225 M^{-1} cm^{-1}). The substrate concentration was 0.2 mM. In order to initiate the ATEE assay, 350 μL of the corresponding enzyme solution (4–7 μM) was added to 650 μL of 0.05 M MES buffer, pH 7.0, at 25 °C. Standard errors in k_{cat} and K_M for all values reported are below 10%.

Active Site Titration. In order to determine the concentration of wild-type and mutant subtilisins, active site titrations were carried out by adding 0, 2, 4, 6, 8, 10, 12, 14, 20, 30, and 50 μL of CI2 (1.84 μM in 0.1 M Tris-HCl, pH 8.6) to 100 μL of a 1:20 diluted subtilisin stock solution, respectively. The buffer used for dilution was 0.1 M Tris-HCl, pH 8.6. All samples were then supplemented with the same buffer to a total volume of 150 μL and incubated for 30 min at room temperature. To determine the activity of the samples, 50 μL of each CI2–subtilisin mixture was added to 950 μL of substrate solution containing 1 mM s-AAPFpNA in 0.1 M Tris-HCl, pH 8.6. The initial velocities were measured by following the absorption at 412 nm. A molar extinction coefficient of 6965 M^{-1} cm^{-1} was used for the determination of the CI2 concentration (S. Jackson, personal communication).

Table I: Activities of Wild-Type and Mutant *B. amyloliquefaciens* Subtilisin BPN′ with s-XAPFpNA Substrates[a,b]

P_4 residue X	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($\times 10^5$ M^{-1} s^{-1})
		Wild-Type Subtilisin	
Ala	35	0.15	2.33
Val	34	0.08	4.25
Ile	30	0.08	3.75
Leu	24	0.06	4.0
Phe	36	0.06	6.0
		Y104F	
Ala	36	0.21	1.71
Val	30	0.10	3.0
Ile	33	0.06	5.5
Leu	16	0.05	3.2
Phe	24	0.04	6.0
		I107V	
Ala	36	0.14	2.57
Val	31	0.06	5.17
Ile	26	0.08	3.25
Leu	21	0.04	5.25
Phe	26	0.06	4.33
		I107A	
Ala	12	0.36	0.33
Val	31	0.16	1.94
Ile	35	0.07	5.0
Leu	34	0.03	11.33
Phe	18	0.04	4.5
		I107G	
Ala	0.32	0.23	0.014
Val	1.5	0.23	0.065
Ile	8.8	0.04	2.2
Leu	27	0.1	2.7
Phe	15	0.05	3.0

[a] Substrate residues s-XAPFpNA correspond to succinyl-X-Ala-Pro-Phe-pNA, where X represents the P_4 substrate residues Ala, Val, Ile, Leu, or Phe. [b] Assays were performed at 25 °C in 0.1 M Tris-HCl, pH 8.6 (see Materials and Methods).

RESULTS

Expression and Purification of Wild-Type and Mutant Subtilisins. All mutant subtilisins were expressed and secreted into the culture medium. Each mutant subtilisin was purified to >95% purity as determined by $NaDODSO_4$–PAGE. The protein yields varied between 6 mg/L for the I107G mutant enzyme to 35 mg/L for the Y104F mutant enzyme.

The *N-trans*-cinnamoylimidazole titration method (Bender et al., 1966) previously used for the determination of active enzyme concentrations depends critically on the ratio of rates of deacylation and acylation and so determines activity of subtilisin mutants only when the active site is not radically altered. We used instead the chymotrypsin inhibitor 2 (CI2) for active site titrations of wild-type and mutant proteins. In order to confirm the accuracy of this method, we also used *N-trans*-cinnamoylimidazole to carry out the active site titration of wild-type subtilisin. Both methods gave the same results within experimental error. The tight binding of CI2 to subtilisin (K_i = 2.9 × 10^{-12} M; Longstaff et al., 1990) yielded sharp titration points for all mutant enzymes, thus facilitating accurate and reproducible determinations of active subtilisin variants.

Steady-State Kinetics Parameters. (a) Wild-Type Subtilisin. The k_{cat} values of wild-type subtilisin (Table I) with substrates carrying the P_4 residues Val, Ile, and Phe (termed Val, Ile, and Phe substrates, respectively) are similar to that of the Ala substrate (35 s^{-1}, Table I). Leu at the P_4 position, however, gives a slightly lower value of k_{cat} of 24 s^{-1}. In contrast, the K_M values follow a clear tendency; as the size of the P_4 side chain increases, the K_M value decreases from

0.15 mM for the Ala substrate to 0.06 mM for the Phe substrate. The value of k_{cat}/K_M varies from 2.33×10^5 M^{-1} s^{-1} to 6.0×10^5 M^{-1} s^{-1} for the same range of substrates. The value of k_{cat}/K_M decreased in the order F > V > L > I > A.

(*b*) *Y104F Mutant Subtilisin.* The value of k_{cat} with Y104F subtilisin is 16 s^{-1} for the Leu substrate, 24 s^{-1} for the Phe substrate, and 36 s^{-1} for the Ala substrate (Table I), i.e., similar to the respective values with wild-type enzyme. Further, the K_M values follow essentially the same trend as observed for wild type; K_M decreases with increasing size of the P_4 side chain. Consequently, the differences in the k_{cat}/K_M values are larger than for the wild-type subtilisin, ranging from 1.71×10^5 M^{-1} s^{-1} for the Ala substrate to 6.0×10^5 M^{-1} s^{-1} for the Phe side chain. The P_4 side-chain preference of the Y104F mutant enzyme is therefore slightly increased toward large hydrophobic side chains compared with wild type. The value of k_{cat}/K_M thus decreases in the order F > I > L > V > A.

(*c*) *I107V, I107A, and I107G Mutant Subtilisins.* The value of k_{cat} for the I107V mutant enzyme varies from 21 s^{-1} for the Leu substrate to 36 s^{-1} for the Ala substrate (Table I). Again, the K_M values show the same general trend as for wild-type and the Y104F mutant subtilisin. K_M varies from 0.14 mM for the Ala substrate to 0.04 mM for the Leu substrate. k_{cat}/K_M values are between 2.57×10^5 M^{-1} s^{-1} and 5.25×10^5 M^{-1} s^{-1} for the Ala substrate and the Leu substrate, respectively. However, the general trend of increased catalytic efficiency toward large side chains at P_4 observed with wild type and the Y104F mutant is not apparent for the I107V mutant, for which the order of k_{cat}/K_M is V = L > F > I > A.

Whereas for the I107A mutant enzyme the k_{cat} values for the Val, Ile, and Leu substrates are similar to that for wild-type subtilisin, the k_{cat} value toward both the Ala and the Phe substrate is decreased 2-fold. The value of K_M of the I107A mutant enzyme decreases with increasing size of the P_4 side chain of the substrate as observed for all other subtilisin variants. However, in comparison to wild-type subtilisin, this effect is more pronounced: the K_M value for the Ala substrate is 2-fold higher and that for the Leu substrate 2-fold lower than for wild-type subtilisin. The highest catalytic efficiency is observed with the Leu substrate (1.13×10^4 M^{-1} s^{-1}). Hence, the order of k_{cat}/K_M compared with wild-type and the Y104F mutant enzyme changes as there is a preference of I107A for Ile and Leu substrates over the Phe substrate. The value of k_{cat}/K_M is in the order L > I > F > V > A.

Of the mutants prepared in this study, I107G exhibits the largest changes in k_{cat} and K_M values. k_{cat} values vary from 0.32 s^{-1} for the Ala substrate to 27 s^{-1} for the Leu substrate. Although the k_{cat} for I107G for the Leu substrate is higher compared with wild-type subtilisin, k_{cat} values for all other substrates are substantially below wild-type level. Both k_{cat} and K_M show marked trends. The k_{cat} value increases with increasing size of the amino acid side chain at the P_4 position of the substrate and the K_M value decreases. The k_{cat}/K_M varies from 1.4×10^3 M^{-1} s^{-1} for the Ala substrate to 3.0×10^5 M^{-1} s^{-1} for the Phe substrate. The I107G mutant enzyme shows, therefore, a 215-fold increase in k_{cat}/K_M for Phe compared with Ala as the P_4 residue. The observed order of k_{cat}/K_M is F > L > I $\gg$ V > A.

(*d*) *Kinetic Parameters of Wild-Type and Mutant Subtilisins with ATEE.* ATEE is a substrate for subtilisin (Schwert & Takenaka, 1955; Philipp & Bender, 1983) and binds in the S_1 pocket. k_{cat}/K_M values of wild-type and mutant enzymes were measured for ATEE as substrate (Table II) in order to measure any effect of the amino acid replacements in the S_4 pocket on the geometry of the catalytic triad and/or the S_1 pocket. The kinetics varies with mutation. Whereas the k_{cat}/K_M value for wild-type subtilisin is 5.07×10^3 M^{-1} s^{-1}, that of the I107V mutant enzyme is 1.4-fold decreased ($k_{cat}/K_M = 3.29 \times 10^3$ M^{-1} s^{-1}) and those of the I107A and Y104F mutant enzymes are both 2.6-fold decreased ($k_{cat}/K_M = 1.96 \times 10^3$ M^{-1} s^{-1} and 1.92×10^3 M^{-1} s^{-1}, respectively). The largest decrease of activity, however, is observed for the I107G mutant subtilisin, which hydrolyzes ATEE with a 15-fold reduced value of k_{cat}/K_M (0.35×10^3 M^{-1} s^{-1}).

Table II: Catalytic Efficiencies of Wild-Type and Mutant Subtilisins BPN′ with Acetyl-L-tyrosine Ethyl Ester[a]

enzyme	k_{cat}/K_M ($\times 10^3$ M^{-1} s^{-1})
wild type	5.07
Y104F	1.92
I107V	3.29
I107A	1.96
I107G	0.35

[a] Assays were performed at 25 °C in 0.05 M MES, pH 7.0 (see Materials and Methods).

(*e*) *Yields of Mutant Enzymes.* Subtilisin is synthesized as a precursor in the form of pre-prosubtilisin (Wells et al., 1983) with the pro sequence required for the correct folding of the enzyme. Since the pro sequence is autocatalytically cleaved from the protease domain (Ikemura & Inouye, 1988), one possible effect of increased S_4 pocket specificity is a reduction in subtilisin expression yields due to insufficient autocatalytic processing (Power et al., 1986). Accordingly, a clear correlation between reduced yields and increase in P_4 substrate residue specificity was observed as yields varied from 6 mg/L cell culture for the mutant with high specificity (I107G) to 17.5 mg/L for mutants with low specificity as the I107V mutant.

DISCUSSION

The ability of subtilisin BPN′ to discriminate between large and small hydrophobic amino acid side chains is increased >200-fold by substituting Ile107 located at the bottom of the S_4 specificity pocket by smaller amino acid residues. To our surprise, we found that mutation of the S_4 site causes not only a change in specificity of long substrates that bind to this site but also a radical drop in the rate of hydrolysis of a short substrate.

The walls of the S_4 pocket are formed by the protein backbone segments 101–104 and 126–128 while the hydrophobic side chains of residues Leu96, Tyr104, Ile107, and Leu126 contribute to the hydrophobic character of the S_4 pocket (Figure 1). The side chains of Leu96 and Leu126 both line the side walls of the S_4 pocket. Tyr104 and Ile107 are situated at the entrance and bottom of the pocket, respectively. Tyr104 has previously been shown to undergo an induced-fit movement, possibly as a consequence of the adjustable size of the S_4 pocket (Robertus et al., 1972a,b; McPhalen & James, 1988; Takeuchi et al., 1991). In the crystal structure of the chymotrypsin inhibitor 2–subtilisin complex (McPhalen et al., 1985), most intermolecular contacts <4 Å to the P_4 residue (Ile56) of CI2 are made by Tyr104 and Ile107, whereas Leu96 and Leu126 also interact with P2 and P3 residues of CI2, respectively. Therefore, the preferred targets for changing the S_4 pocket specificity of subtilisin appear to be residues Tyr104 and Ile107. Our strategy was to improve the specificity of subtilisin for large hydrophobic P_4 side chains by either deleting the hydrogen bond of Tyr104

to Ser130 or increasing the size of the S_4 pocket by systematically deleting methylene groups from the side chain at position 107. Removing the hydrogen bond between residues 104 and 130 was predicted to reduce the binding of small P_4 substrates into the pocket due to an increase in the flexibility of the 104 side chain. However, since the observed effect of the Y104F mutation on the catalytic parameters (Table I) is small, the hydrogen bond, which was thought to lock the Tyr104 side chain into position, may not be as important as initially assumed. Nevertheless, we observed the predicted increase in specificity toward large hydrophobic side chains at the P_4 position.

There is a clear trend in specificity of wild-type, I107V, I107A, and I107G mutant subtilisins for s-XAPFpNA substrates (cf. Table I). The more the S_4 pocket is enlarged by deletions of methyl groups of residue 107, the stronger is the requirement for larger aliphatic P_4 amino acid side chains to reconstitute catalytic efficiency at wild-type level. Previous changes in subtilisin P_1 specificity were calculated by comparing the k_{cat}/K_M values of mutant and wild-type subtilisin for one particular P_1 substrate (Wells et al., 1987). The conventional definition of specificity, however, is the comparison of k_{cat}/K_M values of two (or more) substrates competing for the same enzyme. This is the more useful comparison to use as a working definition of specificity since it mirrors the practical requirements for a specific protease, i.e., the specificity of an engineered protease in distinguishing between different substrates. Using the conventional definition of specificity, the I107G mutant subtilisin has one of the largest improvements in substrate specificity so far reported for subtilisin. Further, the achievement of the improved specificity profile of this mutant enzyme is not at the expense of enzymatic activity.

Changes in P_4 substrate specificity may result from localized changes in P_4–S_4 interactions or structural changes on mutation that emanate from the S_4 pocket to the active site. The rate of hydrolysis of the short substrate ATEE, which binds to the P_1 pocket only, is affected by mutation at the S_4 site (cf. Table II), showing that the latter does occur. The lesion in the I107G mutant subtilisin caused by the changes in the active site geometry is complemented by the presence of an appropiate P_4 amino acid side chain.

Conclusion. Single amino acid substitutions in the S_4 specificity pocket of subtilisin made on the basis of structural information available for subtilisin and the CI2–subtilisin complex can, to some extent, result in rationally designed mutant subtilisins with predicted properties. The unexpected result that mutation of site S_4 decreases the activity of subtilisin toward a short substrate means that we have generated a mutant of subtilisin of unusual specificity, requiring a substrate with a bulky side chain at the P_4 position. Subtilisins containing the mutation Ile107 → Gly, perhaps in combination with other mutations, could well be of use in processes that require a particular specificity: for example, the use of a specific protease for cleavage or as a ligase (Abrahmsen et al., 1991).

ACKNOWLEDGMENT

We thank Drs. M. Bycroft and D. Buck for helpful advice and discussion. We are grateful to T. Gray and T. Buss for synthesizing oligonucleotides and to P. Skelton for performing the mass spectroscopy.

REFERENCES

Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakoff, A., & Wells, J. A. (1991) *Biochemistry 30*, 4151–4159.

Bender, M. L., Begue-Canton, M. L., Blakely, R. L., Brubacher, L. J., Feder, J., Gunter, C. R., Kezdy, F. J., Killheffer, J. V., Marshall, T. H., Miller, C. G., Roeske, R. W., & Stoops, J. K. (1966) *J. Am. Chem. Soc. 88*, 5890–5914.

Bode, W., Papamokos, E., Musil, D., Seemueller, U., & Fritz, H. (1986) *EMBO J. 5*, 813–818.

Bott, R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B., & Power, S. (1988) *J. Biol. Chem. 263*, 7895–7906.

Carter, P., & Wells, J. A. (1987) *Science 237*, 394–399.

Carter, P., Abrahmsen, L., & Wells, J. A. (1991) *Biochemistry 30*, 6142–6148.

Delmar, E. G., Largman, C., Brodrick, J. W., & Geokas, M. C. (1979) *Anal. Biochem. 99*, 316–320.

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.

Grøn, H., Morten, M., & Breddam, K. (1992) *Biochemistry 31*, 6011–6018.

Ikemura, H., & Inouye, M. (1988) *J. Biol. Chem. 263*, 12959–12963.

Kawamura, F., & Doi, R. H. (1984) *J. Bacteriol. 160*, 442–444.

Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331–358.

Longstaff, C., Campbell, A. F., & Fersht, A. R. (1990) *Biochemistry 29*, 7339–7347.

McPhalen, C. A., & James, N. G. (1988) *Biochemistry 27*, 6582–6598.

McPhalen, C. A., Svendsen, I., Jonassen, L., & James, M. N. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7242–7246.

Mitsui, Y., Satow, Y., Watanabe, Y., Hirono, S., & Iitaka, Y. (1979) *Nature 277*, 447–452.

Morihara, K. (1974) *Adv. Enzymol. Relat. Areas Mol. Biol. 41*, 179–243.

Nakajima, K., Powers, J. C., Ashe, B., & Zimmerman, M. (1979) *J. Biol. Chem. 254*, 4027–4032.

Philipp, M., & Bender, M. L. (1983) *Mol. Cell. Biochem. 51*, 5-32.

Power, S. D., Adams, R. M., & Wells, J. A. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3096–3100.

Robertus, J. D., Alden, R. A., Birktoft, J. J., Kraut, J., Powers, J. C., & Wilcox, P. E. (1972a) *Biochemistry 11*, 2439–2449.

Robertus, J. D., Kraut, J., Alden, R. A., & Birktoft, J. J. (1972b) *Biochemistry 11*, 4293–4303.

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.

Schwert, G. W., & Takenaka, Y. (1955) *Biochim. Biophys. Acta 16*, 570–575.

Takeuchi, Y., Noguchi, S., Satow, Y., Kojima, S., Kumagai, I., Miura, K., Nakamura, K. T., & Mitsui, Y. (1991) *Protein Eng. 4*, 501–508.

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.

Thomas, P. G., Russell, A. J., & Fersht, A. R. (1985) *Nature 318*, 375–376.

van der Laan, J. M., Teplyakov, A. V., Kelders, H., Kalk, K. H., Misset, O., Mulleners, L. J. S. M., & Dijkstra, B. W. (1992) *Protein Eng. 5*, 405–411.

Wells, J. A. (1987) in *Protein Engineering*, pp 279–287, Alan R. Liss, Inc., New York.

Wells, J. A., & Estell, A. D. (1988) *Trends Biochem. Sci. 13*, 291–297.

Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A., & Chen, E. Y. (1983) *Nucleic Acids Res. 11*, 7911–7925.

Wells, J. A., Powers, D. B., Bott, R. R., Graycar, T. P., & Estell, D. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1219–1223.

Wright, C. S., Alden, R. A., & Kraut, J. (1969) *Nature 221*, 235–242.