Proc. Natl. Acad. Sci. USA
Vol. 96, pp. 9497–9502, August 1999
Biochemistry

Selection for improved subtiligases by phage display

SHANE ATWELL*† AND JAMES A. WELLS*†‡§¶

*Department of Protein Engineering, Genentech, Inc., 1 DNA Way, South San Francisco, CA 94080; †Graduate Group in Biophysics and ‡Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143; and §Sunesis Pharmaceuticals, Inc., 3696 Haven Avenue, Suite C, Redwood City, CA 94063

Edited by Paul R. Schimmel, The Scripps Research Institute, La Jolla, CA, and approved June 8, 1999 (received for review February 26, 1999)

ABSTRACT **Engineering enzyme activity has been challenging because of uncertainties in structure–function relationships and difficulties in screening a large number of mutant enzymes. A product capture strategy using phage display is presented here for the selection of improved enzymes from a large library of variants (>10^9 independently derived mutants). Subtiligase, a double mutant of subtilisin BPN′ that catalyzes the ligation of peptides, was displayed on phage. Twenty-five active site residues were randomly mutated in groups of four or five to yield six different libraries that were independently sorted. Variants that ligated a biotin peptide onto their own extended N termini were selectively captured. Mutant subtiligases were identified that had increased ligase activity. The selection also yielded unexpected subtiligase mutants having residues known to improve the stability and oxidative resistance of wild-type subtilisin. These studies are exemplary for the use of phage to improve enzyme function when it is closely linked to a selectable catalytic event.**

Enzymes catalyze a variety of biological reactions, making them extremely attractive targets for protein engineering. There are two basic approaches to engineering enzymes. One is a structure-based approach where one generates several designed mutants by modifying specific residues on the enzyme. Another approach is random mutagenesis followed by *in vivo* screening or selection, which offers the tremendous advantage that many variants can be assayed quickly, typically 10^3 to 10^7 (for recent reviews, see refs. 1 and 2). However, in these selections the mutant proteins are tested in a biological setting, which can complicate and limit the range of engineered properties and restrict the possible substrates and reactions. Thus, a general and robust *in vitro* selection method is highly desirable.

Phage display is an extremely powerful *in vitro* selection technique where >10^9 protein or peptide variants can be subjected to selection for improved binding properties (for reviews, see refs. 3 and 4). By selecting for improved binding to transition-state analogs, modest improvements have been obtained for catalytic antibodies displayed on phage (5–8). However, transition-state analog binding does not correlate well with improved catalysis (6, 8). Specialized selections using reactive substrates (9), inhibitors (10), active site ligands (11), and reactive products (12) have also been used in attempts to select for improved catalysts.

A more general functional selection has been developed by investigators interested in novel catalytic RNAs. Catalytic RNA molecules have been efficiently evolved for ligase-type reactions by using a product capture approach wherein the reaction modifies the RNA molecule itself and allows for affinity capture (for review, see ref. 13). Recently, product capture schemes have been used to demonstrate selective enrichment of active staphylococcal nuclease (14), glutathione *S*-transferase (15), and DNA polymerase (16) displayed on phage. However, improved catalysts were not selected.

We wished to test and expand the product capture approach using phage display by sorting large libraries of a mutated enzyme, subtiligase. Subtiligase is a variant of the bacterial serine protease subtilisin BPN′ that was engineered to carry out the ligation of peptides (17). Subtiligase catalyzes the ligation of C-terminal activated peptides to the N-terminal α-amine of peptide and protein acceptors (Fig. 1*A*) (reviewed in ref. 18). We reasoned that catalytic activity might be selected for by requiring that subtiligase on phage catalyze the attachment of a biotin-labeled peptide onto its own N terminus. The biotin-tagged phage could be captured by binding to immobilized neutravidin (Fig. 1*B*). An advantage of testing the product capture scheme on a well-characterized system such as this is that the results can be compared with those from more than a decade of site-directed mutagenesis studies on subtiligase (17, 18) and its parent subtilisin (reviewed in refs. 19, 20).

Here, we report the implementation of the product capture strategy with subtiligase. As expected, the active site residues, including the catalytic triad and oxyanion hole, were completely conserved in the selection. Two new mutants were identified that increased the activity of subtiligase. Interestingly, many other mutants were selected that likely increased stability and resistance to oxidation, as these very same mutants are known to improve these functions in subtilisin. Thus, it appears that product capture and phage display is a powerful approach that directs simultaneous selection for a number of improved properties that are driven by selection for an optimally functional enzyme on phage.

MATERIALS AND METHODS

Synthesis of Peptides. The subtiligase inhibitor biotin-AAF-chloromethylketone (21) was synthesized from sulfo-*N*-hydroxysuccinimide-LC-Biotin (Pierce) and H-AAF-chloromethylketone (Bachem). 2-Aminobenzoyl (Abz)-AAPFglcK-NH_2 (where glc is glycolate) was synthesized according to ref. 18 with fluorenylmethoxycarbonyl chemistry and *N*-Boc-*o*-aminobenzoic acid (Bachem). H-GY(NO_2)-NH_2 [where Y(NO_2) is 3-nitro-L tyrosine] was synthesized from Boc-Gly-OH and Y(NO_2) ethyl ester (Sigma) in dimethylacetamide with hydroxybenzotriazoletetramethyluronium hexafluoro phosphate, followed by ammonolysis (22) and deprotection. The remaining peptides were synthesized by using standard methods (23). All peptides were purified on a preparative C18 HPLC column and lyophilized. Concentrations were calculated from dry weight. Purity was monitored by HPLC and masses verified by mass spectrometry.

Construction of Subtiligase Display Vectors. pSA08541 was constructed by subcloning a fragment of pMal-p2 (New England Biolabs) encoding the *lacI*^{q} gene, the P_{tac} promoter, and the signal peptide from maltose binding protein (malsignal) between the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office.

Abbreviations: ABz, 2-aminobenzoyl; glc, glycolate; Y(NO_2), 3-nitro-L-tyrosine.

¶To whom reprint requests should be addressed. E-mail: jaw@sunesis-pharma.com.

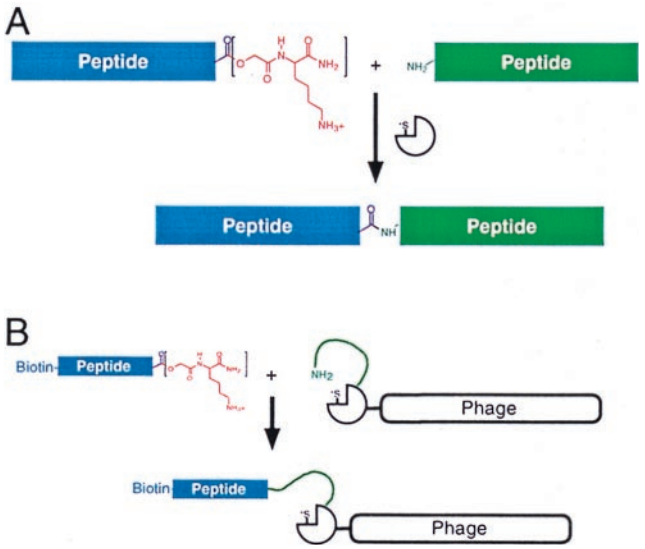

FIG. 1. (*A*) Ligation reaction catalyzed by subtiligase. (*B*) Scheme for selecting active subtiligase mutants on phage by requiring that the enzyme attach a biotin-labeled peptide onto its extended N terminus. The biotin-labeled subtiligase phage are then captured with immobilized neutravidin.

*Eco*RI and NsiI sites of pH0753 (24). The prosubtiligase gene was subcloned from pLA1019 (see below) into this vector between the *Nsi*I and *Xba*I sites to generate pSA0906, a subtiligase display vector expressing malsignal$^{1-26}$-S-pro$^{1-77}$-subtiligase$^{1-275}$-V-*-gIIIp$^{249-406}$ (where * indicates an amber stop suppressed in XL1-Blue cells, and gIIIp is the gene III protein of bacteriophage M13). Derivatives of pSA0906 lacking the pro domain (Δpro) (pSA0928), the calcium-binding loop residues 75 to 83 (ΔCa) (pSA0933), or both (Δpro/ΔCa) (pSA0940) were constructed by site-directed mutagenesis (25). Further derivatives with the pro region coexpressed were constructed by subcloning the signal peptide and pro domain from pSA0906 into the *Nhe*I site of pSA0928 and pSA0940 to generate pSA0953 (Δpro + pro) and pSA0947 (Δpro/ΔCa + pro).

In the course of developing the subtiligase display system, a number of other constructs were tested. It was found that the P_{tac} promoter produced higher display levels of subtiligase than the P_{phoA} promoter; subtilisin displayed more than subtiligase; the gIIIp anchor produced much higher display levels than the gVIIIp anchor; and the stII secretion sequence was equivalent to the mal signal (S.A. and J.W., unpublished results).

Subtiligase display vectors with various N-terminal extension lengths were constructed by site-directed mutagenesis of ΔCa, with the following five inserts between the pro domain and the second residue of ΔCa subtiligase: GLGY, GLSEGGY, GLSEGGGSGY, GLSEGGGSEGGGY, GLSEGGGSEGGGSEGY.

For the experiments determining whether ligation was intra- or intermolecular, we used Δpro/ΔCa and clone 46 of the Δpro/ΔCa extension library (pSA0940.46) (see below) to avoid autoprocessing artifacts. A pSA0940.46 derivative with the catalytic Cys-221 mutated to Ala (pSA0988) was generated by site-directed mutagenesis.

Phage Manipulations. Phage manipulations were as described (26). Phage were captured for ELISAs by using a polyclonal antisubtilisin antibody or neutravidin (Pierce). To capture biotinylated subtiligase phage with neutravidin, phage in PBS containing 1 mM DTT were incubated with biotin-inhibitor or biotin substrate. Nonspecific capture of the phage became evident above 1 μM inhibitor, so we used a low concentration of the inhibitor. Selection of N-terminal extension variants and product capture ELISAs used 100 nM biotin-inhibitor or 100 nM biotin-substrate (in PBS, 10% DMSO) for 30 min, followed by 10-fold dilution with casein blocking buffer (Pierce) and capture with neutravidin in one well. The six active site libraries were sorted separately by reacting with 1 μM substrate (in 1% DMSO) for 30 min followed by polyethylene glycol precipitation of phage. Phage were resuspended in 0.4 ml of blocking buffer and aliquots distributed equally among eight wells of a microtiter plate precoated with neutravidin to capture the biotinylated phage. After washing, bound phage were propagated by addition of midlog phase *Escherichia coli* XL1-Blue (Stratagene) directly to wells.

Increases in display levels were determined by fitting antibody capture ELISA curves to a + (b-a)/(1 + (([phage]/c)^d)*((b - 0.5)/(0.5 - a))), in which a, b, c, and d are fitting parameters. C is the $EC_{0.5}$, the concentration of phage needed to give an OD_{492} of 0.5. The $EC_{0.5}$ of the wild type was divided by the $EC_{0.5}$ of the mutant ($EC_{0.5}$wt/$EC_{0.5}$mut) to give the relative display level. Ligation activities were determined by using the $EC_{0.5}$wt/$EC_{0.5}$mut determined from a product capture ELISA and by normalizing for display levels by dividing by the $EC_{0.5}$wt/$EC_{0.5}$mut from the antibody capture ELISA. In cases where the ELISA signal was weak, comparisons were made with the $EC_{0.1}$, the concentration of phage needed to generate an ELISA signal of 0.1.

Construction of Phage Libraries. The N-terminal extension library was made by inserting the sequence XXX(SEGGG)$_2$XX, where X is any of the 20 amino acids, between the signal peptide and the first residue of ΔCa subtiligase by using degenerate oligonucleotides and loop-out mutagenesis on the Δpro/ΔCa template. The mutagenized DNA was transformed into *E. coli* strain SS320 by electroporation (refs. 25 and 27; D. Sidhu and J.W., unpublished work). The total number of transformants was 4.0×10^9.

Active site libraries were designed with the aid of a graphics workstation (Silicon Graphics, Mountain View, CA), by using the coordinates for subtilisin BPN′ with an inhibitor [2st1.pdb (28)] or the pro domain bound [1spb.pdb (29)]. The six libraries were constructed by first replacing each set of target codons in pSA0940.46 with stop codons and then mutagenizing these codons with oligonucleotides containing NNS at these positions (where n = A, C, G, or T and S = C or G) (26). The total number of transformants for the six libraries ranged from 2.3×10^9 to 4.0×10^9. Because the number of transformants for each library exceeded the maximum nucleotide diversity by at least 100-fold, every amino acid combination was probably present.

Expression and Purification of Subtiligase Mutants. Subtiligase mutants expressed poorly in *E. coli*, and proteolytic degradation was evident when expressing ΔCa mutants in *Bacillus subtilis*. Therefore, subtiligase mutants were subcloned into derivatives of pSS5, a *B. subtilis* expression vector, encoding pre-pro subtiligase. This subcloning restored the pro domain and the calcium-binding loop (17). Expression and purification of the soluble variant proteins were as previously described (17), except that PD10 columns (Amersham Pharmacia) were used to exchange the final buffer to 100 mM Tricine (pH 8.0), 5 mM $CaCl_2$. Subtiligase concentrations were determined by densitometric scanning of Coomassie-stained SDS gels by using a known concentration of subtiligase as a standard (30). Most of the mutants that were characterized as soluble proteins expressed as well as the parent subtiligase.‖ However, the A152C variant did not express well and so could not be characterized as a soluble protein. The yield of the M124L/L126V variant was 3-fold greater than the parent subtiligase, whereas the Y217C/N218T was 8-fold less, the M222F 11-fold less, and M222S/S224A 2-fold less than the parent.

Rates of Ligation Measured by HPLC and Fluorescence Assays. Aliquots (22.5 μl) of the ligation reaction: 300 μM

‖Amino acids are indicated by single- or triple-letter code; mutations are designated by the single-letter amino acid code for the wild-type residue followed by its sequence position and the mutant residue. N- and C-terminal peptide substituents are indicated by chemical names hyphenated before or after the single-letter peptide sequence.

biotin-SGGKAAPFglcK-NH_2/3 mM H-GLRY-NH_2/100 mM Tricine (pH 8.0)/0.005% Tween 20/1 mM DTT/1 μM enzyme were quenched at 15-sec intervals with 2.5 μl acetic acid and stored on ice. Samples were analyzed on a C18 HPLC column by using a 0–90% acetonitrile/0.1% trifluoroacetic acid gradient, monitoring product peak areas at 214 nm.

Fluorescence quenching of ABz-AAPFglcK-NH_2 by ligation to H-GY(NO_2)-NH_2 (100 μM ABz-AAPFglcK-NH_2/1 mM H-GY(NO_2)-NH_2/0.005% Tween 20/100 mM Tricine (pH 8.0)/1 mM DTT/5 mM $CaCl_2$/5 μM enzyme) was monitored on an SLM 8000 fluorimeter (SLM—Aminco, Urbana, IL) by excitation at 333 nm and monitoring emission at 400 nm after normalizing the signal and subtracting the baseline. Because of low yields of expressed protein, M222F and Y217C/N218T fluorescence kinetics were compared with wild type at 1 μM enzyme.

RESULTS

Phage Display of Functional Subtiligase. The first challenge in developing a product capture selection for subtiligase on phage was to demonstrate functional display for the enzyme that would ligate peptides onto its own N terminus. By way of background, subtilisin BPN′ is naturally expressed as a pre-pro protein. The presequence is removed on secretion. The pro domain participates in folding the enzyme and is subsequently cleaved in an autocatalytic process. Finally, the enzyme binds calcium, which is important to its thermal stability. Deletion of the calcium-binding loop destabilizes the enzyme but allows it to fold reversibly and independently of the pro domain (31).

A phage ELISA was used to test for the display of subtiligase on phage. Phage from serial dilutions of concentrated stocks ($\approx 10^{13}$ phage particles per ml) were captured by using an immobilized antisubtilisin antibody. The bound phage were detected with an antiphage antibody conjugated with horseradish peroxidase. Using the phage ELISA, we could just detect display of subtiligase on phage when it was expressed under the control of a P_{tac} promoter as the pre-pro protein fused to gIIIp. The display of subtiligase was improved 350-fold by removing the calcium loop (residues 75 to 83; ΔCa) and allowing subtiligase to fold more reversibly (Fig. 2). Coexpressing the pro domain (Δpro + pro) had little effect, but removing the pro domain altogether (Δpro) abolished display. Combining the Δpro and ΔCa mutation still permitted reasonable display levels and had the advantage of allowing the enzyme to fold to its mature form without the assistance of the N-terminal pro domain and without the requirement that subtiligase autoproteolyze the pro domain. We therefore chose the Δpro/ΔCa construct for the selection experiments.

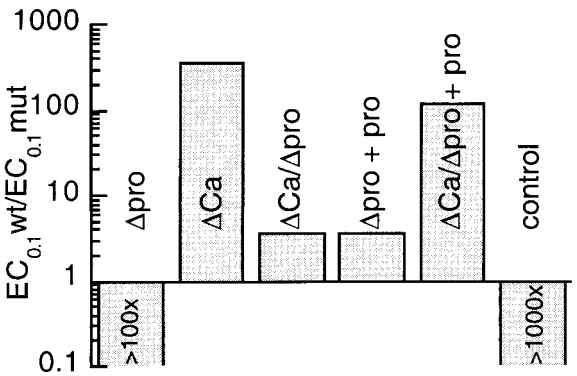

FIG. 2. Improved display of engineered subtiligase. Serially diluted phage were captured with an antisubtilisin antibody in a phage ELISA. The bars show the ratio of wild-type to mutant phage needed to generate the same ELISA signal and thus indicate the improvement in phage display for the mutant compared with the wild type.

To demonstrate that the displayed subtiligase retained enzyme activity, we analyzed its ability to react with a covalent serine protease inhibitor, biotin-AAF-chloromethylketone, which is known to react with the active site cysteine of thiol-subtilisin (S221C) (21). We found that subtiligase phage that were reacted with the biotin-labeled inhibitor were efficiently captured on neutravidin plates in proportion to their display on phage (data not shown). Incubating the phage with 1 mM DTT to reduce the active site cysteine lead to a 100-fold increase in capture by the inhibitor (data not shown). We hypothesize that DTT reduces a mixed disulfide bond between the active site cysteine and a small molecule thiol from the culture media. Thus, the selections, ELISAs, and kinetic experiments were performed after reduction with DTT. In no case did the addition of DTT interfere with the ligation reaction.

Subtiligase Phage Catalyzed Self Ligation with a Biotin-Labeled Substrate. The strategy for selecting active enzymes required that the N terminus reach the active site to function as a nucleophile. The N terminus of mature subtilisin is ≈30 Å from the P1′ recognition site (28, 29). Structural considerations and the behavior of unstructured peptides (32) suggested that the N terminus needed to be extended by 10 to 15 residues to reach the active site. We therefore constructed several derivatives of Δpro/ΔCa subtiligase with increasing N-terminal extension lengths. There was a dramatic improvement in the ligation efficiency once the N terminus was extended beyond six residues (Fig. 3*A*). The effective concentration of the N terminus near the active site is probably in the range of 10 to 100 mM (32), and thus it is reasonable to expect the reaction to be intramolecular. Nonetheless, it was important to demonstrate directly that ligation was occurring intramolecularly. This was done by showing that subtiligase phage lacking an N-terminal extension do not ligate a peptide onto a catalytically inactive subtiligase containing the N-terminal extension (Fig. 3*B*).

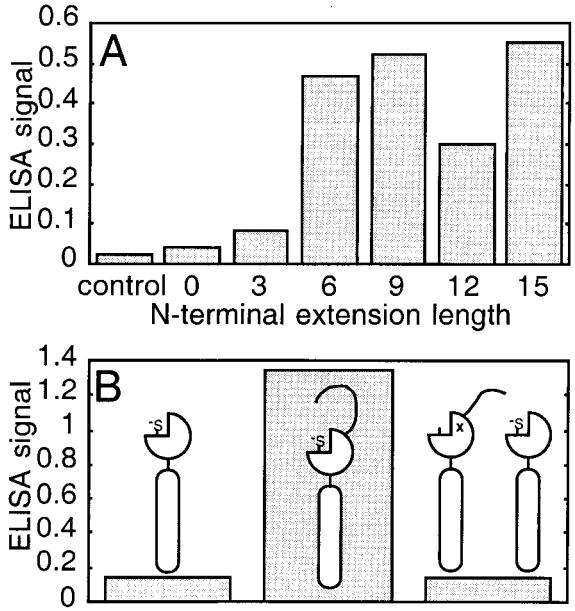

FIG. 3. (*A*) Capture of subtiligase containing various N-terminal extensions. Concentrated phage (10 μl; approximately 10^{13} ml^{-1}) were incubated with 100 nM substrate (biotin-SGGKAAPFglcK-NH_2 in PBS, 10% DMSO). After the samples were reacted, the mixtures were diluted 10-fold with blocking buffer, captured with neutravidin, and ELISA signal measured. The ELISA signals are not normalized for phage concentrations or display levels so that the drop in activity for the 12-mer is probably not significant. (*B*) Intramolecular vs. intermolecular ligation. The ligation reaction was measured for subtiligase phage, subtiligase phage with a 12-residue N-terminal extension, and subtiligase phage mixed with inactive subtiligase phage containing the N-terminal extension.

Previous studies have shown that efficient ligation depends on the sequence of the nucleophilic peptide (33). Thus it was desirable to optimize the residues of the N-terminal nucleophile on subtiligase. We constructed a subtiligase-phage library with a 15-residue extension having the form $XXX(SEGGG)_2XX$. This library had three randomized residues at the N terminus, followed by a 10-residue flexible linker and two more randomized residues joined to the natural N terminus of subtiligase. To avoid folding and autoprocessing artifacts, the Δpro/ΔCa construct was used, and the extension was inserted between the signal sequence and the natural N terminus. The library of N-terminally extended subtiligases was reacted with biotin-labeled peptide substrate, and the phage that became labeled with biotin were captured on neutravidin plates. The extent of biotin phage capture was measured by ELISA. After three rounds of sorting, there was little further increase in captured phage; thus, we assumed the functional selection was largely complete (data not shown; ref. 27). Sixteen clones were assayed for ligase activity and sequenced (Table 1). Glycine and alanine residues, which are preferred at P1′ for ligation (33), were commonly found in the selection.

Three similar N-terminal extension libraries were constructed based on the full-length parent subtiligase, the ΔCa, and the Δpro/ΔCa + pro variant subtiligases. Sorting results from these libraries demonstrated a strong P1′ preference for glycine (27). Among the four libraries, there was a prevalence of amber stop codons, rare AGG codons, positively charged residues, and whole codon and frameshift mutations that nevertheless yielded good display levels (for review, see ref. 4). These were probably selected to reduce the host toxicity of subtiligase by reducing protein expression. We had noticed that *E. coli* bearing the subtiligase phagemid grow slower and are more prone to lysis (data not shown). Mutations that decrease subtiligase toxicity and therefore increase culture density and phage production, while still yielding the minute amounts of protein needed for incorporation into the phage coat, probably lead to an overall increase in phage displaying subtiligase.

We chose the clone with the N-terminal GLR sequence because it gave an ELISA signal 2-fold better than the designed 15-residue extension, and its P1′ and P2′ matched the known subsite preferences for subtiligase in solution experiments (34). The Δpro/ΔCa + pro variant subtiligase having an optimized N-terminal extension satisfied our criteria of a subtiligase that can catalyze intramolecular ligation and obviates the need for pro domain assisted folding and autoproteolysis.

Selection for Improved Catalysis and Characterization of Mutant Subtiligases. Subtiligase is a designed double mutant of subtilisin BPN′. The S221C mutation leads to a 300-fold preference for aminolysis vs. hydrolysis (35); the P225A variant further increases the rate of ligation 10-fold (17).

To further optimize the subtiligase active site for ligation, we targeted the 25 residues nearest the catalytic site, including Asp-32, His-64, Asn-155, and Pro-225 (Fig. 4). Four to five contiguous residues were randomized in six libraries: residues 31 to 34, 64 to 68, 124 to 127, 152 to 155, 217 to 220, and 222 to 225. The libraries were sorted separately for five to seven rounds of selection for their ability to self ligate the biotin-labeled substrate until the yield of product-captured phage no longer increased. At this point, ten to twenty clones from each library were assayed and sequenced (Table 2). We then analyzed soluble forms of the eight most active mutants and the most frequently isolated mutant, I31L/S33T. The mutations were subcloned into the parent subtiligase, which contains the calcium-binding loop and pro domain and lacks the N-terminal extension. These soluble subtiligase variants were expressed in *B. subtilis* (see *Materials and Methods*).

Two methods were used to measure the ligation reaction in solution. First the appearance of product was monitored with HPLC by using the same substrate as that used for the selection (300 μM biotin-SGGKAAPFglcK-NH_2) and a tetrapeptide with the same first three residues as the N-terminal extension

Table 1. N-terminal extension library selectants

	Signal-XXX$(SEGGG)_2$XX-ligaseΔCa					
Clone	X1	X2	X3	X14	X15	ELISA
45	A	*	T	T	E	2.0
42	R	K	E	S	P	1.5
43	G	Q	A	R	N	1.2
41 (2)	N	K	R	E	*	1.1 ± 0.1
46	G	L	R	D	R	1.0
37	*	T	S	N	M	0.9
44	*	D	R	R	P	0.6
39	*	L	V	G	N	0.6
52	A	V	G	K	Q	0.6
54	A	D	*	G	G	0.5
51	G	G	P	L	S	0.5
40	G	S	R	H	D	0.5
47	R	G	A	Q	Q	0.3
53	Q	K	S	Y	R	0.3
50	S	A	M	V	L	0.2

Mutations identified in the first three residues in the N terminus of subtiligase and after three rounds of selection. The product capture ELISA signal is indicated in the far-right column to provide an estimate of the extent of display. Two of the clones (#41) were determined to be identical after sequencing. * indicates an amber stop, which is suppressed with Gln in XL1-blue cells.

of the subtiligase (3 mM H-GLRY-NH_2, 1 μM enzyme). The acceleration of the reaction with each mutant compared with the reaction with the wild-type enzyme is shown in Fig. 5.

It was difficult to detect the ligation product at lower concentrations by HPLC. To solve this problem, a fluorescence donor-quencher pair of substrates was synthesized to more sensitively monitor the ligation reaction in a fluorescence quenching assay (36). *O*-aminobenzoyl-AAPFglcK-NH_2 is excited at 320 nm and fluoresces at 407 nm; however, the ligation product, aminobenzoyl-AAPFGY(NO_2)-NH_2, is 95% quenched, which facilitated direct and sensitive measurements of ligation rates. The initial velocity of quenching was measured for the wild-type enzyme and various mutants (40 μM ABz-AAPFglcK-NH_2/1 mM H-GY(NO_2)-NH_2/5 μM enzyme). There is reasonable agreement between the HPLC and fluorometric assays (Fig. 5). Because the fluorescence assay used a substrate concentration of 40 μM, which is far below the K_M for the activated peptide (about 1 mM; ref. 17), this assay provides a better indication of the

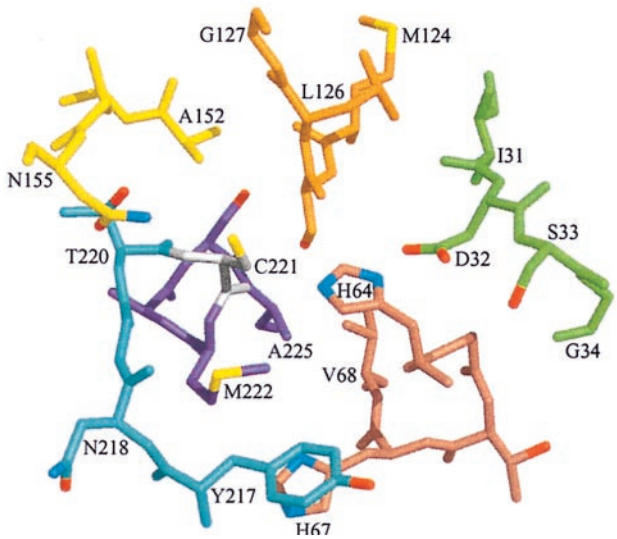

FIG. 4. Structure of the active site region of subtiligase. Residues that were randomized are indicated by distinct colors for each of the six libraries (green for 31–34; pink for 64–68; orange for 124–127; yellow for 152–155; blue for 217–220; and purple for 222–225). Heteroatoms in the catalytic triad are colored by atom type.

Table 2. Active site library selectants

Clone	Residues randomized					Increased display	Increased activity
Library 1	I31	D32	S3	G34			
1.1 (17/4)	L	.	T	.		11	1.0
Library 2	H64	G65	T66	H67	V68		
2.1 (5/1)	.	.	.	Y	.	1.5	0.13
2.2 (4/3)	.	.	.	A	.	3.2 ± 0.2	0.13 ± 0.003
2.3	.	.	.	*	.	2.1	0.065
Library 3	M124	S125	L126	G127			
3.1	L	.	G	.			
3.2	L	.	V	.		1.1	2.3
3.3	L	A	.	.		0.53	2.3
3.4	L	G	A	.		0.69	0.41
3.5	L	G	C	.		2.9	1.2
3.6	.	G	C	.		3.0	1.0
3.7	I	A	V	.		2.7	0.48
3.8	I	.	.	.			
Library 4	A152	A153	G154	N155			
4.1 (14/1)	C	.	.	.		3.1	0.068
4.2 (2/1)	S	.	.	.		0.40	0.079
Library 5	Y217	N218	G219	T220			
5.1	.	S	.	.		2.6	0.63
5.2	A	S	.	.		3.4	0.72
5.3	C	T	.	.		3.4	1.9
5.4	V	T	.	.		1.7	0.58
5.5	M	A	.	.		2.7	0.19
5.6	G	H	.	.		2.2	0.61
Library 6	M222	A223	S224	A225			
6.1 (2/2)	G	.	.	.		0.49	1.9
6.2	S	.	.	.		0.74	1.9
6.3	S	.	A	.		0.49	1.7
6.4	A	.	.	G		0.46	1.4
6.5	V	.	.	.		0.29	0.19
6.6	*	.	.	.		1.5	0.14
6.7	F	.	.	.		0.17	2.7
6.8	.	S	C	.		0.31	0.79

Mutations identified in subtiligase (see Fig. 4) after random mutation of six regions and sorting for four to seven rounds under product capture conditions as in Table 1. Periods indicate conservation of parent subtiligase residues. Asterisks indicate amber stops, suppressed as Gln in XL1-blue cells. Numbers in parentheses indicate the number of instances found for a particular amino acid sequence followed by the number of different DNA sequences. The improved activity on phage is normalized for improved display (see *Materials and Methods*). Assays on clones with the same amino acid sequence indicate that the results are reproducible to ±15%.

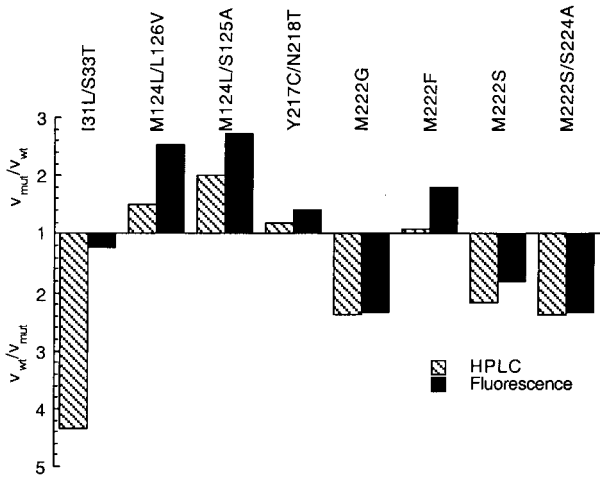

FIG. 5. Kinetics of mutants measured by HPLC and fluorescence quenching. Increases in activity vs. wild type are plotted above the axis, and decreases vs. wild type are plotted below. In the HPLC determination (300 μM biotin-SGGKAAPFglcK-NH_2/3 mM H-GLRY-NH_2/1 μM enzyme), the initial velocity of product formation catalyzed by the parent subtiligase, $v_{w.t.}$, was 58 $\mu M\ s^{-1}$. In the fluorescence assay (40 μM Abz-AAPFglcK-NH_2/1 mM H-GY(NO_2)-NH_2/5 μM enzyme), $v_{w.t.} = 0.14\ \mu M\ s^{-1}$. Rates were reproducible to ±10%.

change in k_{cat}/K_M (37, 38). From these data, it was clear that M124L/S125A and M124L/L126V are more than 2-fold improved over the parent subtiligase.

There was a poorer correlation between the increased activity measured on phage (Table 2) and the increased activity of the purified proteins (Fig. 5). The poor correlation could be caused by the different concentrations of substrates used, the presence of the calcium loop in the soluble enzymes, or other differences in assay conditions.

DISCUSSION

Here we developed a product capture method for selecting subtiligase variants displayed on phage. Despite the fact that the selections were performed on subtiligase, the selection results largely mirror mutagenic studies of its parent subtilisin BPN′ (19). For example, 15 of the 25 residues randomized in subtiligase were strongly or completely conserved as the wild type (Table 2) (39). Asp-32 and His-64 of the catalytic triad were completely conserved, as were Asn-155 and Thr-220 of the oxyanion hole (19, 40).

Interestingly, even though the selection was intended to improve ligase activity, many of the selected subtiligase mutations are known to improve the thermodynamic or oxidative stability for wild-type subtilisin. For instance, N218S dramatically stabilizes subtilisin to heat denaturation (31). In our subtiligase selection, Asn-218 was mutated largely to Thr or Ser. Met-222 is

a known site of oxidative inactivation in subtilisin. Though Met-222 is conserved in subtilisins, in the subtiligase selection it was mostly mutated to small amino acids that are known to be the least deleterious to subtilisin activity (41) (Table 2). Oxidation of Met-222 on phage may also account for the selection of Tyr or Ala in place of the nearby His-67, also a highly conserved residue in the subtilisin family; model peptide studies show that proximal histidines can accelerate the oxidation of methionine (42). Although we have not directly measured the oxidative stability or thermal stability of these subtiligase variants, it is probable that the selected residues function in subtiligase as they do in subtilisin. Subtilisin has been a very robust scaffold for transferring functions between homologs (19) and between subtilisin and subtiligase (17, 33).

Given the difference in enzyme mechanism between subtilisin and subtiligase, it is not surprising that some of the selected residues are specific to subtiligase vs. subtilisin. Subtiligase was originally created as a double mutant: the S221C mutation promotes ligation over hydrolysis, and P225A enhances activity by contracting the supporting helix and providing more room for the larger cysteine nucleophile (17). Here we show that when residue 225 was fully randomized and selected on phage, the P225A was largely conserved, validating the original design.

Other residues that appear to function differently in subtiligase vs. subtilisin include Ser-125. Ser-125 is conserved throughout the subtilisin superfamily (39) and forms a hydrogen bond with the main-chain nitrogen of residue 221. Nonetheless, Ser-125 is not conserved in the subtiligase selections. This mutation may allow repositioning of cysteine nucleophile in subtiligase. Leu-126 is conserved among subtilisins and is important for subtilisin activity (43), but was not conserved in the subtiligase selection.

It is interesting that the two mutations that are beneficial to subtiligase but harmful to subtilisin, L126V and M222F (41), are on opposite sides of Cys-221 (Fig. 4). It is possible that these mutations, like P225A, reposition the active-site Cys-221 relative to the substrate. Further structural analysis is needed to test these hypotheses.

Finally, two mutants were isolated from the 124–127 library that were more active than the original subtiligase: M124L/L126V was 2.5-fold improved, and M124L/S125A was 2.7-fold improved. M124L might have been selected for oxidation resistance; however, it is more likely that it was selected for its direct effect on catalysis because Met-124 oxidation has little effect on subtilisin activity.

Subtilisin does not naturally contain disulfides, but designed disulfide bonds are known to form readily (ref. 44 and references therein). It is interesting that in some cases disulfides may have been selected here in subtiligase. A152C was found in 14 of 16 selectants. This residue is close to the active-site Cys-221 (5.0 Å Cα to Cα). The fact that treating subtiligase phage with DTT increases the yield of captured phage 100-fold suggests that the active-site thiol is substantially blocked in the culture media. Formation of a disulfide between Cys-221 and Cys-152 could provide a selective advantage by protecting Cys-221 from reversible or irreversible oxidation. Several other nearby cysteines appeared in the selection, including L126C and S224C (6.3 Å and 5.2 Å from Cys-221). Besides protection from oxidation, formation of disulfides between these residues and Cys-221 might have contributed to folding or stability, because most of the mutants displayed better on phage (Table 2). Under our selection conditions *in vitro*, these disulfide bonds would be reduced by DTT to generate the free Cys-221 and therefore an active enzyme.

Some mutants appeared to be selected to increase display on phage. I31L/S33T was completely conserved in the 31–34 library but does not have improved activity (Fig. 5). This double mutant displays 10-fold better on phage than the parent subtiligase, giving it a strong selective advantage (Table 2).

In summary, the product capture selection on subtiligase reflected many functional features of subtilisin, such as the importance of certain residues for catalysis, stability, and resistance to oxidation. Several subtiligase mutants were selected for improved activity and a number of others for improving functional enzyme display. Future selections should provide additional variants that, when combined, may allow for more dramatic improvements in subtiligase activity. It should also be possible to focus selection on particular properties besides catalysis. If it were desirable to improve the stability properties, one could alter the library construction and selection conditions to include denaturants or oxidants; the phage themselves are quite tolerant. Considerable effort was necessary to display active subtiligase on phage. Nonetheless, the selections appear to be very robust, and they offer the advantage that conditions can be modulated *in vitro* and that massive numbers of enzyme variants ($>10^9$) can be screened simultaneously. We believe this could be a very powerful approach to engineering catalytic and other functional properties in enzymes.

We thank the oligonucleotide synthesis group, Michelle Arkin, Andrea Cochran, Greg Weiss, and Cliff Quan for helpful advice, the Bioorganic Chemistry Department at Genentech, Inc. for reagents, equipment, and helpful advice, Andrew Braisted for biotin-SGGKAAPFglcK-NH_2, and Abraham de Vos for subtiligase coordinates. S.A. was supported by Genentech, Inc.

1. Kast, P. & Hilvert, D. (1997) *Curr. Opin. Struct. Biol.* **7,** 470–479.
2. Zhao, H. & Arnold, F. H. (1997) *Curr. Opin. Struct. Biol.* **7,** 480–485.
3. Clackson, T. & Wells, J. A. (1994) *Trends Biotechnol.* **12,** 173–184.
4. Wilson, D. R. & Finlay, B. B. (1998) *Can. J. Microbiol.* **44,** 313–329.
5. Hansson, L. O., Widersten, M. & Mannervik, B. (1997) *Biochemistry* **36,** 11252–11260.
6. Baca, M., Scanlan, T. S., Stephenson, R. C. & Wells, J. A. (1997) *Proc. Natl. Acad. Sci, USA* **94,** 10063–10068.
7. Fujii, I., Fukuyama, S., Iwabuchi, Y. & Tanimura, R. (1998) *Nat. Biotechnol.* **16,** 463–467.
8. Arkin, M. R. & Wells, J. A. (1998) *J. Mol. Biol.* **284**, 1083–1094.
9. Janda, K. D., Lo, C., Li, T., Barbas, C. F., Wirsching, P. & Lerner, R. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2532–2536.
10. Light, J. & Lerner, R. A. (1995) *Bioorg. Med. Chem.* **3,** 955–967.
11. Widersten, M. & Mannervik, B. (1995) *J. Mol. Biol.* **250,** 115–122.
12. Janda, K. D., Lo, L.-C., Lo, C.-H., Sim, M.-M., Wang, R., Wong, C.-H. & Lerner, R. A. (1997) *Science* **275,** 945–948.
13. Zhang, B. & Cech, T. R. (1998) *Chem. Biol.* **5,** 539–553.
14. Pedersen, H., Holder, S., Sutherlin, D. P., Schwitter, U., King, D. S. & Schultz, P. G. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10523–10528.
15. Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L, Neri, P., Winder, G. & Neri, D. (1999) *J. Mol. Biol.* **286,** 617–633.
16. Jestin, J-L, Kristensen, P. & Winter, G. (1999) *Angew. Chem. Int. Ed. Engl*. **38,** 1124–1127.
17. Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakoff, A. & Wells, J. A. (1991) *Biochemistry* **30,** 4151–4159.
18. Braisted, A. C., Judice, J. K. & Wells, J. A. (1997) *Methods Enzymol.* **289,** 298–313.
19. Wells, J. A. & Estell, D. A. (1988) *Trends Biochem. Sci.* **13,** 291–297.
20. Perona, J. J. & Craik, C. S. (1995) *Protein Sci.* **3,** 337–360.
21. Philipp, M. & Bender M. L. (1983) *Mol. Cell. Biochem.* **51,** 5–32.
22. Bodanszky, M. & Bodanszky, A. (1994) in *The Practice of Peptide Synthesis* (Springer, New York), 2nd Ed. rev., p. 165.
23. Barany, G. & Merrifield, R. B. (1980) in *The Peptides*, eds. Gross, E. & Meinenhoffer, J. (Academic, New York), Vol. 2, pp. 1–284.
24. Atwell, S., Ultsch, M., de Vos, A. M. & Wells, J. A. (1997) *Science* **278,** 1125–1128.
25. Kunkel, T. A., Bebenek, K. & McClary, J. (1991) *Methods Enzymol.* **204,** 125–139.
26. Lowman, H. B. (1998) in *Methods in Molecular Biology: Combinatorial Peptide Library Protocols,* ed. Cabilly, S. (Humana, Totowa), Vol. 87, pp. 249–264.
27. Atwell, S. (1999) Ph. D. thesis (University of California, San Francisco).
28. Bott, R., Ultsch, M., Kossiakoff, A., Graycar, T., Katz, B. & Power, S. (1988) *J. Biol. Chem.* **263,** 7895–7906.
29. Gallagher, T., Gilliland, G., Wang, L. & Bryan, P. (1995) *Structure (London)* **3,** 907–914.
30. Cunningham, B. C., Jhurani, P., Ng, P. & Wells, J. A. (1989) *Science* **243,** 1330–1336.
31. Brian, P., Alexander, P., Strausberg, S., Schwarz, F., Lan, W., Gilliland, G. & Gallagher, D. T. (1992) *Biochemistry* **31,** 4937–4945.
32. Creighton, T. E. (1993) *Proteins: Structures & Molecular Properties* (Freeman, New York), pp. 171–199.
33. Chang, T. K., Jackson, D. Y., Burnier, J. P. & Wells, J. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 12544–12548.
34. Gron, H., Meldal, M. & Breddam, K. (1992) *Biochemistry* **31,** 6011–6018.
35. Natatsuka, T., Sasaki, T. & Kaiser, E. T. (1987) *J. Am. Chem. Soc.* **109,** 3808–3810.
36. Meldal, M. & Breddam, K. (1991) *Anal. Biochem.* **195,** 141–147.
37. Fersht, A. (1985) *Enzyme Structure and Mechanism* (Freeman, New York), pp. 101–103, 118–119, 200–205.
38. Bender, M. L., Clement, G. E., Gunter, C. R. & Kezdy, F. J. (1964) *J. Am. Chem. Soc.* **86,** 3697–3703.
39. Siezen, R. J. & Leunissen, J. A. M. (1997) *Protein Sci.* **6,** 501–523.
40. Braxton, S. & Wells, J. A. (1991) *J. Biol. Chem.* **266,** 11797–11800.
41. Estell, D. A., Graycar, T. P. & Wells, J. A. (1985) *J. Biol. Chem.* **260,** 6518–6521.
42. Schoneich, C., Zhao, F., Wilson, G. S. & Borchardt, T. T. (1993) *Biochim. Biophys. Acta* **1158,** 307–322.
43. Rheinnecker, M., Eder, J., Pandey, P. S. & Fersht, A. R. (1994) *Biochemistry* **33,** 221–225.
44. Mitchinson, C. & Wells, J. A. (1989) *Biochemistry* **28,** 4807–4815.