32. T. Godefroy-Colburn, M. J. Gagey, A. Berna, C. Stussi-Garaud, *J. Gen. Virol.* **67**, 2233 (1986).
33. C. Stussi-Garaud, J. C. Garaud, A. Berna, T. Godefroy-Colburn, *ibid.* **68**, 1779 (1987).
34. L. Sequeira, *Trends Biotechnol.* **2**, 25 (1984).
35. Y. Watanabe, I. Ooshika, T. Meshi, Y. Okada, *Virology* **152**, 414 (1986).
36. C. Gigot, M. Kopp, C. Schmitt, R.G. Milne, *Protoplasma* **84**, 31 (1975).
37. R. N. Beachy *et al.*, *EMBO J.* **4**, 3047 (1985); Z. L. Chen, M. A. Schuler, R. N. Beachy, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8560 (1986); D. E. Rochester, J. A. Winter, D. M. Shah, *EMBO J.* **5**, 451 (1986).
38. U. K. Laemmli, *Nature (London)* **227**, 680 (1970).
39. R. S. Nelson, P. Powell-Abel, R. N. Beachy, *Virology* **158**, 126 (1987).
40. R. T. Fraley *et al.*, *Biotechnology* **3**, 629 (1985).
41. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
42. A. O. Jackson and B. A. Larkins, *Plant Physiol.* **57**, 5 (1976).
43. M. M. Bradford, *Anal. Biochem.* **72**, 248 (1976).
44. H. Towbin, T. Staelin, J. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350 (1979); J. Symington, M. Green, K. Brackmann, *ibid.* **78**, 177 (1981).
45. We thank E. Anderson and P. Dube for skilled technical assistance, M. Dyer and D. Droste for growing plants, C. Holt for affinity purifying the antiserum to the 30-kD protein, and N. Burkhart for secretarial assistance. Supported by a grant from the Monsanto Company.

20 April 1987; accepted 27 June 1987

Engineering Enzyme Specificity by "Substrate-Assisted Catalysis"

PAUL CARTER AND JAMES A. WELLS

A novel approach to engineering enzyme specificity is presented in which a catalytic group from an enzyme is first removed by site-directed mutagenesis causing inactivation. Activity is then partially restored by substrates containing the missing catalytic functional group. Replacement of the catalytic His^{64} with Ala in the *Bacillus amyloliquefaciens* subtilisin gene (the mutant is designated His64Ala) by site-directed mutagenesis reduces the catalytic efficiency (k_{cat}/K_m) by a factor of a million when assayed with *N*-succinyl-L-Phe-L-Ala-L-Ala-L-Phe-*p*-nitroanilide (sFAAF-pNA). Model building studies showed that a His side chain at the P2 position of a substrate bound at the active site of subtilisin could be virtually superimposed on the catalytic His side chain of this serine protease. Accordingly, the His64Ala mutant hydrolyzes a His P2 substrate (sFAHF-pNA) up to 400 times faster than a homologous Ala P2 or Gln P2 substrate (sFAAF-pNA or sFAQF-pNA) at *p*H 8.0. In contrast, the wild-type enzyme hydrolyzes these three substrates with similar catalytic efficiencies. Additional data from substrate-dependent *p*H profiles and hydrolysis of large polypeptides indicate that the His64Ala mutant enzyme can recover partially the function of the lost catalytic histidine from a His P2 side chain on the substrate. Such "substrate-assisted catalysis" provides a new basis for engineering enzymes with very narrow and potentially useful substrate specificities. These studies also suggest a possible functional intermediate in the evolution of the catalytic triad of serine proteases.

RATIONAL DESIGN OF ENZYME SPECIFICITY BY PROTEIN engineering should expand the range and utility of biological catalysts. Site-directed mutagenesis techniques (*1*) have already been applied to modify the substrate specificity of subtilisin (*2–4*), trypsin (*5*), tyrosyl-tRNA (transfer RNA) synthetase (*6*), carboxypeptidase (*7*), alcohol dehydrogenase (*8*), and aspartate aminotransferase (*9*). These studies have focused on altering specificity by changing residues that make direct contact with the substrate. Here we present an alternative approach, termed "substrate-assisted catalysis," in which part of the catalytic machinery of the enzyme is removed and appropriately supplied by a similar functionality from a bound substrate. In this way substrates are distinguished primarily by their ability to actively participate in the catalytic mechanism permitting the design of extremely specific enzymes.

We have chosen subtilisin, a serine class endopeptidase, as a model to test the concept of substrate-assisted catalysis. In the hydrolysis of peptide bonds by subtilisin, His^{64} acts both as a catalytic base and acid in the formation of the acyl-enzyme intermediate and in a similar fashion in the subsequent deacylation step (*10*). A model of a substrate containing a histidine residue at the P2 position (*11*) bound to subtilisin (Fig. 1) indicated that the Nδ1 and Nε2 nitrogens can almost be superimposed (to about an angstrom unit) on the corresponding nitrogens of the catalytic histidine (His^{64}). This suggested that if the histidine in the catalytic triad of subtilisin was replaced by an alanine by means of site-directed mutagenesis, a histidine from the substrate might supply the missing catalytic functional group.

Construction and expression of His64Ala mutant subtilisin. Maturation of the primary subtilisin gene product (preprosubtilisin) to subtilisin in *Bacillus subtilis* is believed to be mediated by autoproteolysis that involves trace amounts of active subtilisin (*12*). The His64Ala mutation (*13–19*) caused a severe reduction in secretion of mature subtilisin that was presumed to result from a large reduction in catalytic activity (*20*). However, it was possible to process and subsequently purify the weakly active His64Ala enzyme by co-culturing *B. subtilis* cells harboring the His64Ala mutant gene with *B. subtilis* cells carrying an active subtilisin gene (helper) (*21–24*).

Stringent precautions were taken to ensure the purification of His64Ala subtilisin away from helper subtilisin and any other contaminating proteases. First, the mutant subtilisin was expressed in the *B. subtilis* host BG2036 (*19*), which is deficient in chromosomal copies of the genes for alkaline protease (subtilisin) and neutral protease. To minimize helper contamination, the ratio of helper cells to His64Ala cells in the fermentation culture was adjusted to 1:1000. A functionally silent Ser24Cys mutation that is located on the surface of subtilisin (*17*) was introduced into the

The authors are in the Department of Biomolecular Chemistry, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.

Table 1. Kinetic analysis of mutant subtilisins with the substrates, *N*-succinyl-L-Phe-L-Ala-L-X-L-Phe-*p*-nitroanilide, when X is Ala, Gln, or His. Six hydrolysis assays were performed simultaneously against corresponding substrate blanks in 0.10*M* tris-HCl (*p*H 8.6), 10 m*M* dithiothreitol at 25° ± 0.1°C with a Kontron Uvikon 860 spectrophotometer. Initial reaction rates were determined from the increase in absorbance caused by the release of *p*-nitroaniline [$\epsilon_{410\ nm} = 8480 M^{-1}\ cm^{-1}$ (*34*)] and fitted by linear regression to an Eadie-Hofstee plot to calculate V_{max} and K_m. The k_{cat} values were calculated from V_{max}; the spectrophotometrically determined enzyme concentration was used (*24*). Enzyme concentrations in the assays were about 50 μg/ml for Ser24Cys:His64Ala enzyme and 1 μg/ml for the Ser24Cys enzyme. Standard errors in all determinations were below 20 percent. Slight variation in the absolute kinetic values has been observed between batches of enzyme, but the relative values among substrates have remained constant.

Substrate P2 residue	Ser24Cys			Ser24Cys:His64Ala		
	k_{cat} (sec^{-1})	K_m (μ*M*)	k_{cat}/K_m ($sec^{-1}\ M^{-1}$)	k_{cat} (sec^{-1})	K_m (μ*M*)	k_{cat}/K_m ($sec^{-1}\ M^{-1}$)
Ala	8.1	10	8.0×10^5	8.1×10^{-6}	32	0.25
Gln	7.0	39	1.8×10^5	3.0×10^{-5}	150	0.20
His	4.6	23	2.0×10^5	1.6×10^{-2}	380	42

His64Ala mutant. This accessible cysteine permitted reversible attachment to an activated thiol Sepharose column that enabled purification of the His64Ala mutant away from the cysteine-free helper subtilisin. Finally, the helper subtilisin contained a functionally silent Ala48Glu mutation (*22*) that altered its electrophoretic mobility relative to the Ser24Cys:His64Ala double-mutant subtilisin on native and sodium dodecyl sulfate (SDS)–polyacrylamide gels (*20*). After purification, the Ser24Cys:His64Ala mutant was judged to be more than 99 percent pure by silver-stained SDS (*25*) and native polyacrylamide gel electrophoresis (*20*).

Kinetic analysis of Ser24Cys:His64Ala. The kinetic parameters of Ser24Cys and Ser24Cys:His64Ala were determined for the substrates *N*-succinyl-L-Phe-L-Ala-L-[X]-L-Phe-*p*-nitroanilide (abbreviated sFAXF-pNA), where X is Ala, Gln, or His (Table 1). Wild-type and Ser24Cys subtilisins have essentially identical kinetic parameters for these substrates (*26*) indicating that the Ser24Cys mutation is kinetically silent. By comparison, the His64Ala mutation causes a millionfold decrease in the ratio of k_{cat} to K_m for the Ala and Gln P2 substrates. Almost all of the decrease in catalytic efficiency is caused by a decreased k_{cat} value (up to 10^6 times), although significant increases appear in K_m (up to 4 times). The relatively small changes in K_m may reflect changes in the enzyme-substrate dissociation constant (K_s) although it may also reflect a shift in the rate determining step of the reaction (*27*); we have not determined whether acylation is rate limiting for the His64Ala mutant, as it is for the wild-type enzyme (*28*). In any case, the catalytic histidine contributes a factor of about 10^6 to the enzymatic rate enhancement (Table 1).

Unlike wild-type or Ser24Cys subtilisin, the Ser24Cys:His64Ala enzyme was completely resistant to inhibition by the active site reagent, phenylmethylsulfonyl fluoride (PMSF). This suggests that the catalytic histidine is critical for stable sulfonylation by PMSF. Although the proportion of functional active sites in Ser24Cys:His64Ala enzyme preparations could not be determined directly by such active site labeling, enzyme that was purified further by native gel electrophoresis (*20*) had identical kinetic parameters to Ser24Cys:His64Ala (Table 1) purified as described (*21*).

For the Ser24Cys enzyme, the Ala P2 substrate is preferred over the Gln and His P2 substrates by about four times k_{cat}/K_m (Table 1). Most of this difference is the result of larger K_m values for the Gln and His P2 substrates. For the Ser24Cys:His64Ala enzyme, the relative differences in k_{cat}/K_m toward the Ala and Gln P2 substrates are similar to Ser24Cys. In contrast, the Ser24Cys:His64Ala enzyme hydrolyzes the His P2 substrate more efficiently by factors of 170 and 210 times k_{cat}/K_m, respectively. Essentially all of the increase in k_{cat}/K_m for the His compared to the Ala and Gln P2 substrates results from the k_{cat} term being larger by a factor of 2000 and 500, respectively. The larger K_m values for the His and Gln P2 substrates compared to Ala may reflect reduced binding affinity. Apparently, the catalytic advantage of the His P2 side chain is only realized in the context of the catalytic groups provided by the Ser24Cys:His64Ala enzyme, because the nonenzymatic hydrolysis rates for the His and non-His substrates are essentially the same. Thus, the drop in k_{cat}/K_m caused by the His64Ala mutation is partially restored when cleaving a His P2 substrate. The net effect is a marked increase in substrate preference for a His P2 side chain brought about primarily at the level of catalysis rather than by binding.

The *p*H dependence of amide bond hydrolysis by Ser24Cys:His64Ala. The *p*H dependence of k_{cat}/K_m for wild-type subtilisin shows a sigmoidal increase from *p*H 6 to 8 (*29*) that reflects the titration of the catalytic His^{64} ($pK_a = 7.1 \pm 0.1$). The wild-type *p*H profile remains relatively flat over the range of *p*H 8 to 10 and declines thereafter (*30*).

The *p*H dependence of k_{cat}/K_m is markedly different for the

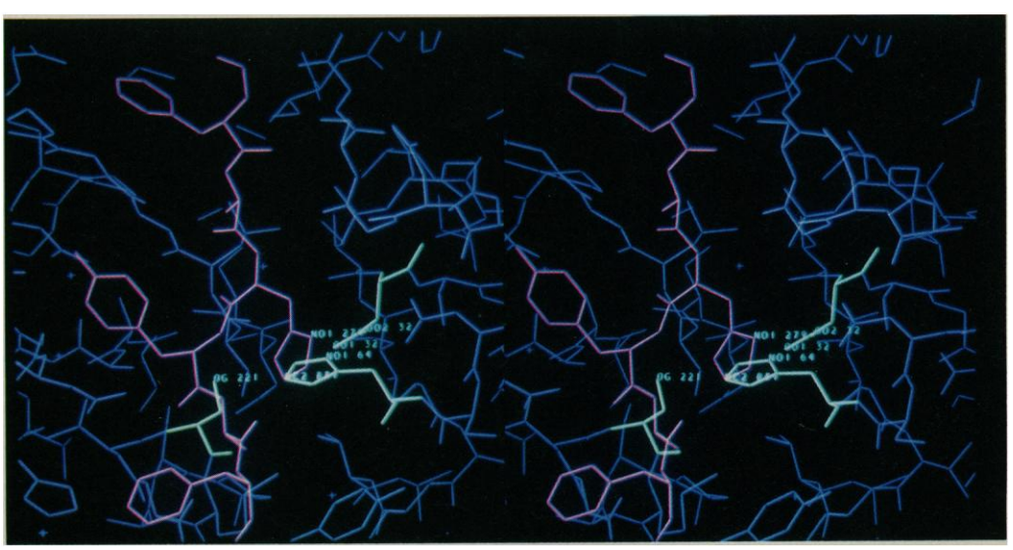

Fig. 1. Stereo view of *B. amyloliquefaciens* subtilisin (blue) containing a modeled bound peptide substrate (pink) having the sequence L-Phe-L-Ala-L-His-L-Tyr-L-Gly-L-Phe representing residues P4 to P2′ of the substrate (*11*). For model building we used the program FRODO (*36*) and an Evans and Sutherland PS300 graphics system. The model was based upon a 2.0 Å x-ray crystallographic study of the product complex (NH_2-L-Phe-L-Ala-L-Ala-Leu-COOH) bound to subtilisin (*37*). The Ala P2 side chain was replaced by a His side chain fixing the main chain position and allowing the His side chain dihedral angles to vary (Table 3). The catalytic triad (Asp^{32}, His^{64}, and Ser^{221}) is shown in green with the His P2 side chain from the substrate superimposed upon the catalytic His^{64} (model corresponds to model S1 in Table 3).

Table 2. Digestion of peptide substrates by Ser24Cys:His64Ala subtilisin. Various synthetic peptides (200 μg) shown were digested with Ser24Cys:His64Ala subtilisn (10 μg) in 20 m*M* tris-HCl (*p*H 8.0), 10 m*M* DTT, 5 percent (v/v) dimethyl sulfoxide, 1 m*M* PMSF (250 μl, total volume) for 20 hours at 37°C. Digestion products were analyzed by reversed phase HPLC as described in the legend to Fig. 3. Digestion products recovered by HPLC were hydrolyzed for 24 hours in 6*N* HCl, 1 percent (v/v) phenol before amino acid analysis, with norleucine as an internal standard. The Cys residues in bovine insulin A and B chains were oxidized to cysteic acid. Sequences are designated by the single letter amino acid code, and arrows indicate potential His P2 cleavage sites. ACTH, adrenocorticotropin hormone; BOP is an open reading frame from the complementary strand to the gene for bacterial rhodopsin of *Halobacterium*. Single letter abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

Peptide source	Sequence	Clevage peptides with Ser24Cys:His64Ala
Inhibin β chain (residues 61–80)	TVIN**H**↓YRMRG**H**↓SPFANLKSC	TVIN**H**Y + RMRG**H**SPFANLKSC
ACTH (residues 1–10)	SYSME**H**↓FRWG	SYSME**H**F + RWG
Ubiquitin (residues 62–76)	CKESTL**H**↓LVLRLRGG-NH_2	Not cleaved
BOP gene product (residues 68–86)	GYE**H**↓FENLRRRAASFQGKY	Not cleaved
Bovine insulin B chain (oxidized)	FVNQ**H**↓LCGS**H**↓LVEALYLVCGERGFFYTPKA	Not cleaved
Bovine insulin A chain (oxidized)	GIVEQCCASVCSLYQLENYCN	Not cleaved

Ser24Cys:His64Ala enzyme. For the sFAAF-pNA substrate, there is an increase of 15 times in k_{cat}/K_m between *p*H 8 and 10 (Fig. 2A). The k_{cat}/K_m ratio shows a linear dependence on hydroxide ion concentration (Fig. 2B) suggesting that a hydroxide ion can act as a catalytic base in the absence of a catalytic histidine side chain. If we extrapolate from the increase in k_{cat}/K_m as a function of hydroxide concentration (2×10^4 sec^{-1} M^{-2}), to the k_{cat}/K_m for Ser24Cys with the same Ala P2 substrate (8×10^5 sec^{-1} M^{-1}), then the equivalent concentration of the hydroxide ion would be about 40*M*.

In contrast, k_{cat}/K_m for hydrolysis of the sFAHF-pNA by Ser24Cys:His64Ala shows a sigmoidal *p*H dependence between *p*H 6 and 8 (Fig. 2C) that is similar to wild-type subtilisin. The pK_a of the activity-dependent group is 6.8 ± 0.1, and almost all of the *p*H-dependent changes in k_{cat}/K_m result from changes in k_{cat}. For the sFAHF-pNA substrate, there is not a strong linear increase in k_{cat}/K_m with hydroxide above *p*H 8 as observed for hydrolysis of sFAAF-pNA. These data therefore suggest that the His P2 substrate side chain or hydroxide ion can substitute functionally for the missing catalytic His^{64}.

The data presented in Table 1 (measured at *p*H 8.6) underestimate the substrate preference for His over Ala (and Gln) because the k_{cat}/K_m for the sFAHF-pNA is maximal at *p*H 8.0 (Fig. 2C), whereas for the sFAAF-pNA substrate it is significantly lower at *p*H 8.0 (Fig. 2A). Thus, for Ser24Cys:His64Ala at *p*H 8.0, we estimate the substrate preference is up to ~400 times for the His P2 substrate over the corresponding Ala or Gln substrates.

Many lines of evidence indicate that the activity we attribute to the Ser24Cys:His64Ala enzyme is not the result of any other protease contamination. First, the extreme substrate preference for His at the P2 position is unlike wild-type subtilisin or any known *Bacillus* protease. Furthermore, the mutant has K_m values that are significantly different from wild-type subtilisin, suggesting differences in substrate binding or catalysis. In addition, the mutant is completely resistant to inhibition by PMSF, unlike other serine proteases. In fact, the kinetic determinations and polypeptide digestion experiments with the Ser24Cys:His64Ala mutant (Table 1 and Fig. 2) were routinely made in the presence of PMSF to exclude any possibility of active helper subtilisin. Moreover, the substrate-dependent *p*H profiles are unlike any protease we are aware of. In addition, preparations of Ser24Cys:His64Ala are extremely pure on the basis of analysis by SDS and native gels (>99 percent). Finally, the kinetic values determined for Ser24Cys:His64Ala that was further purified by native gel electrophoresis (*20*) are essentially the same as those reported in Table 1.

Hydrolysis of polypeptide substrates by Ser24Cys:His64Ala. To evaluate further the specificity of Ser24Cys:His64Ala in comparison with Ser24Cys, we incubated both enzymes with a 20-residue fragment of the inhibin β chain at *p*H 8.0. The choice of the peptide was based on the presence of two histidines (position 5 and 11) along with 16 other amino acids including various large hydrophobic amino acids that are preferred amino acids at the P1 position of wild-type subtilisin (*2*). The peptide (Fig. 3A, peak a) at ~120-fold molar excess over Ser24Cys:His64Ala was cleaved into only two

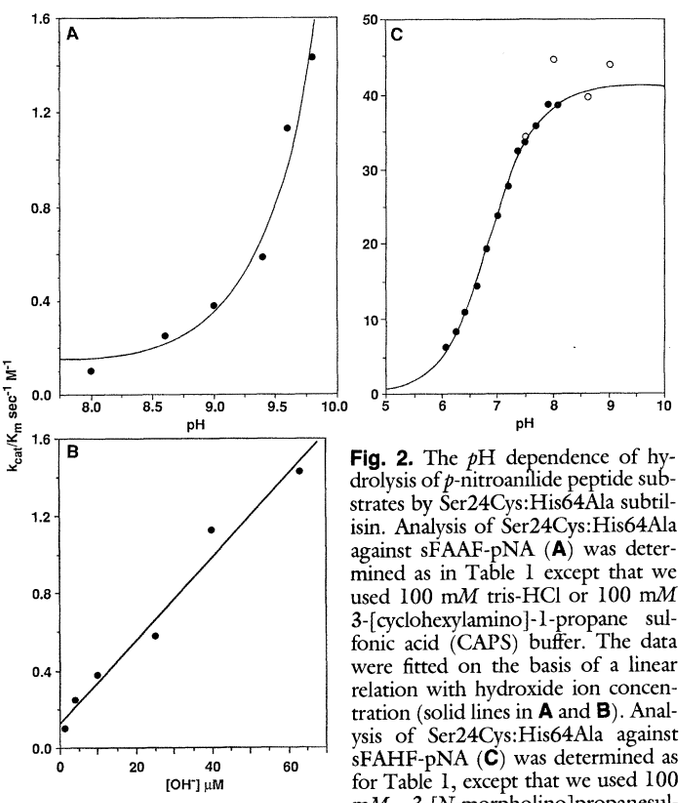

Fig. 2. The *p*H dependence of hydrolysis of *p*-nitroanilide peptide substrates by Ser24Cys:His64Ala subtilisin. Analysis of Ser24Cys:His64Ala against sFAAF-pNA (**A**) was determined as in Table 1 except that we used 100 m*M* tris-HCl or 100 m*M* 3-[cyclohexylamino]-1-propane sulfonic acid (CAPS) buffer. The data were fitted on the basis of a linear relation with hydroxide ion concentration (solid lines in **A** and **B**). Analysis of Ser24Cys:His64Ala against sFAHF-pNA (**C**) was determined as for Table 1, except that we used 100 m*M* 3-[*N*-morpholino]propanesulfonic acid (MOPS) buffer (closed circles) or 100 m*M* tris-HCl (open circles) and then normalized the ionic strength with KCl. The data were fitted to a sigmoid curve (solid line) with the use of a least-squares fit procedure.

pieces (Fig. 3A, peaks b and c) to more than 95 percent after incubation for 2 hours. Amino acid composition analysis of these two peptide fragments indicated cleavage had occurred between Tyr^6 and Arg^7, as would be expected for substrate-assisted catalysis by His^5 located at the P2 position relative to the cleavage site. After a longer (ten times) period of digestion (20 hours) a minor third peak appeared (Fig. 3A, peak X). Analysis showed it to have the same composition as the undigested inhibin peptide. This minor product also appeared in a nonenzymatic blank incubation. No digestion was observed at the second histidine site.

In contrast to the two fragments produced by Ser24Cys:His64Ala, the Ser24Cys enzyme produced at least seven fragments (Fig. 3B) at a similar extent of digestion of starting material (compare 5-minute digestion with Ser24Cys to 2-hour digestion with Ser24Cys:His64Ala). Although none of these seven fragments was sequenced, the first two produced eluted from the high-performance liquid chromatogram (HPLC) at the same positions as peaks b and c in Fig. 3A. Digestion to 95 percent completion of the starting peptide by Ser24Cys (Fig. 3C, 30-minute incubation) produced more than ten different peptide fragments.

Digestion experiments for this and five other peptides are summarized in Table 2. A ten-residue fragment of human adrenocorticotropic hormone (ACTH) was quantitatively cleaved at a single site by Ser24Cys:His64Ala. Analysis of the amino acid composition of the two digestion products confirmed that the cleavage had occurred with a His residue at the P2 position of the substrate as expected. However, digestion of this peptide with Ser24Cys also gave specific cleavage at the same position. This probably resulted because Phe provides a very favorable P1 residue, and the two short peptides liberated are relatively poor substrates for subtilisin. The four other peptides tested (three containing His and one which did not) were not cleaved by Ser24Cys:His64Ala but were cleaved at several sites by the Ser24Cys mutant.

These experiments establish much about the activity, specificity, and utility of the Ser24Cys:His64Ala mutant. In addition to *p*-nitroanilide substrates, the enzyme can cleave normal peptide bonds. Unlike the Ser24Cys enzyme, the specificity of Ser24Cys:His64Ala appears to be limited to sites containing a histidine side chain located at the P2 position of the cleavage site. Furthermore, additional specificity determinants are required because not all His P2 sites are cleaved. We believe that this reflects the normal specificity determinants in the wild-type enzyme, which extend at least from P4 to P2′ (*30*). Peptide substrates were chosen to have His followed by a large hydrophobic amino acid, which is preferred for the P1 site in subtilisin (*2, 30*). The potential cleavage sites for ACTH and the BOP (Table 2) peptide are very similar, except that the uncleaved BOP site contains a Glu at the P1′ position. Little is known about P1′ specificity, but the absence of cleavage at the other His sites may reflect the presence of a Val, a Pro, or a negatively charged residue (Glu or cysteic acid), all of which are very poorly hydrolyzed P1 amino acids (*2, 31*).

Proteolytic instability of heterologous gene products (particularly small proteins) sometimes necessitates their expression as larger fusion polypeptides and recovery by chemical or enzymatic cleavage. The Ser24Cys:His64Ala mutant, or variant thereof, may be especially useful in site-specific proteolysis of fusion polypeptides. First, the enzyme is extremely specific for a P2 histidine compared to its closest isosteric homolog, glutamine. Furthermore, histidine is the third least abundant amino acid in proteins having an average abundance of 2.1 percent in proteins (*32*). In addition, specificity determinants at the other subsites in subtilisin further restrict the cleavage specificity among possible histidine sites. If necessary, the specificity of the P1 binding site in subtilisin can be further restricted by protein engineering (*2–4*). Moreover, subtilisin is unusually stable in denaturants often necessary for solubilization of fusion proteins. Finally, the surface thiol (Ser24Cys) permits easy immobilization or separation.

Substrate-assisted catalysis in design and evolution of serine proteases. The fact that the catalytic efficiency of the Ser24Cys:His64Ala enzyme for the His P2 substrate is 5000 times less than that of the wild type suggests that the His provided by the substrate functions poorly in catalysis. Indeed, the model of the His P2 side chain does not exactly match the catalytic His^{64} (Fig. 1 and Table 3). When the distances for the His P2 and His^{64} imidazoyl nitrogens are matched to the catalytic Ser and Asp (Table 3, model S1), dihedral angles for the His P2 side chain differ from ideality (that is, $\chi_1 = \pm 60°, -180°$; $\chi_2 = \pm 90°$) and the corresponding imidazoyl nitrogens from His^{64} and His P2 are separated by more than an angstrom. In addition, the planes of the histidine side chains from the enzyme and substrate are nearly perpendicular to each other

Table 3. Pertinent bond angles and distances modeled for substrate-assisted catalysis by a His P2 side chain in subtilisin or trypsin as described in Figs. 1 and 4, respectively. Dihedral angles for the His side chains are defined by χ_1 (N–Cα–Cβ–Cγ) and χ_2 (Cα–Cβ–Cγ–Cδ). The hydrogen bond angles [Oγ(Ser)–Hδ(Ser)–Nε2(His)] were calculated from the measured Cβ(Ser)–Oγ(Ser)–Nε2(His) angle, the Nε2(His)–Oγ(Ser) distance and the known Oγ(Ser)–Hδ(Ser) distance (0.96 Å) and the Cβ(Ser)–Oγ(Ser)–Hδ(Ser) bond angle (108.5°) (*35*). The hydrogen bond angles Nδ1(His)–Hε(His)–Oδ2(Asp) were measured after poisitioning Hε(His) in the structure using the known distances: Cγ(His)–Nε2(His) (1.39 Å), Cε(His)–Nε2(His) (1.34 Å) and Nε2(His)–Hε(His) (1.01 Å) and the known angles Cε(His)–Nε2(His)–Hε(His) (125.35°), Cγ(His)–Nε2(His)–Hε(His) (126.35°) (*35*). Hydrogen bond distances were measured from the catalytic Ser (Oγ) and Asp (Oδ1 and Oδ2) to the Nε2 and Nδ1, respectively, of the enzyme His or the substrate His P2. The distances are given between the enzyme His and the modeled substrate His P2 Nε2 and Nδ1 nitrogens. Model 1 (shown in Figs. 1 and 4 for subtilisin and trypsin, respectively) has the His P2 side chain optimized for hydrogen bond distances between the imidazoyl nitrogens, Nε2 and Nδ1, to the catalytic Ser and Asp, respectively. Model 2 has idealized χ angles for the His P2 side chain. Molecular dynamics and energy minimization procedures were not applied to any of these models.

	Angles				Distances (Å)				
	Dihedral		H bonds		Nε2 (His)→ Oγ(Ser)	Nδ1 (His)→		Catalytic His→His P2	
	χ_1	χ_2	(Ser→His)	(His→Asp)		Oδ1(Asp) or Oδ2(Asp)		Nε2/Nε2	Nδ1/Nδ1
Subtilisin									
Catalytic His^{64} (actual)	−167°	85°	148°	156°	3.17	3.34	2.72		
His P2 side chain									
Model S1	−164°	−50°	149°	98°	3.17	3.55	2.72	1.39	1.35
Model S2	−180°	−90°	144°	84°	3.25	3.59	3.34	0.37	1.57
Trypsin									
Catalytic His^{57} (actual)	71°	85°	170°	129°	2.70	3.25	2.70		
His P2 side chain									
Model T1	−155°	−79°	179°	83°	2.78	4.78	3.28	0.98	2.10
Model T2	−180°	−90°	158°	122°	2.48	5.09	3.76	0.58	2.09

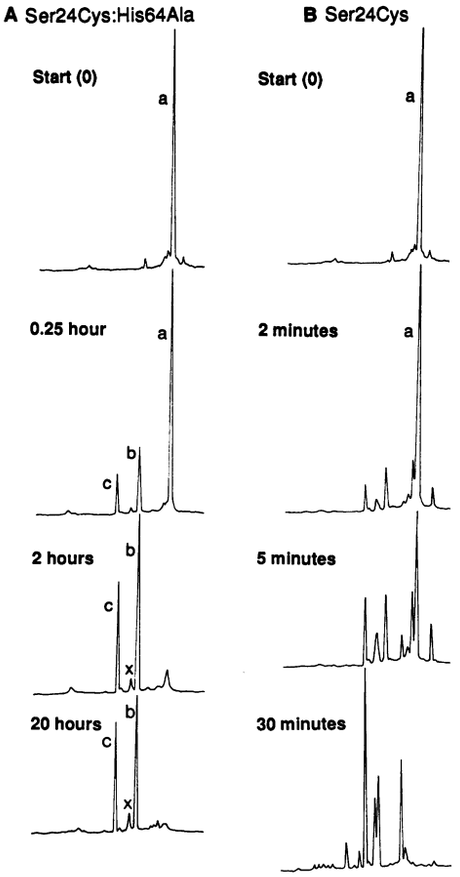

Fig. 3. Hydrolysis of a polypeptide substrate (inhibin β chain) by subtilisin mutant Ser24Cys:His64Ala or Ser24Cys. The synthetic peptide substrate: TVINHYRMRGHSPFANLKSC (100 μg) was digested with 10 μg of the Ser24Cys:His64Ala enzyme (**A**) or with 0.13 μg of Ser24Cys enzyme (**B**). Reaction mixtures (total volume of 250 μl) contained 20 m*M* tris-HCl (*p*H 8.0), 10 m*M* dithiothreitol, 5 percent (v/v) dimethyl sulfoxide, and 1 m*M* PMSF (Ser24Cys:His64Ala only). After indicated times at 37°C, digestion products (monitored at 214 mm) were eluted from a reversed phase HPLC column (Waters, C18) with a gradient (from left to right) of 0 to 50 percent (v/v) acetonitrile in 0.1 percent (v/v) trifluoroacetic acid. Peaks a, b, c in (A) refer to the starting peptide, fragment TVINHY and fragment RMRGHSPFANLKSC, respectively. Peak X has a composition identical to the starting peptide.

(Fig. 1). This model does not appear to have a detrimental effect on the hydrogen bond angle between the Ser and His (Table 3) or lead to bad steric contacts. However, the modeled hydrogen bond angle between His and Asp is less favorable. Further mutagenesis experiments are under way to determine whether Asp^{32} has a significant catalytic role in the His64Ala enzyme. When the dihedral angles are constrained to ideality, only $\chi_1 = 180°$ and $\chi_2 = -90°$ give approximate superposition of His^{64} and His P2 (Table 3, model S2). The hydrogen-bond distances, dihedral angles, and steric contacts are much less favorable when the substrate imidazoyl ring is rotated 180° (opposite to that of His^{64} or model S1 in Table 3) to pair the Nδ1 with Ser^{221} and Nε2 with Asp^{32}. X-ray crystallographic studies will be necessary to clarify the precise mode of binding for the His P2 substrate. Further mutagenesis studies are directed toward producing more efficient substrate-assisted catalysis by a His P2 side chain.

The strategy for altering substrate specificity by changing substrate (or transition-state) binding contacts (*2–9*) has general applicability in that contacts to any part of the substrate can be engineered to change their binding properties. The substrate-assisted catalysis approach is particularly applicable where removal of a catalytic group (such as a nucleophile, or an acidic or basic group) from an enzyme can be precisely replaced by a similar functional group from a bound substrate. Thus, enzymes (such as proteases) that react with substrates having potential catalytic functional groups may be especially good candidates for this approach to engineering specificity.

There are at least two distinct families of serine proteases, the "subtilisin-like" and "trypsin-like," which evolved the catalytic triad from two entirely separate genes (*10*). It is highly unlikely that the precise positioning of the catalytic triad (Ser-His-Asp) evolved in one step but more likely involved intermediates. Model building from a complex between bovine pancreatic trypsin inhibitor (BPTI) and bovine trypsin shows a His side chain substituted at the equivalent P2 position (Cys^{14}) of BPTI can be reasonably superimposed upon the catalytic histidine (Fig. 4). Moreover, the dihedral angles, hydrogen-bond angle, and relevant distances between the Nδ1 and Nε2 imidazoyl nitrogens from the enzyme or P2 substrate histidines to the catalytic Asp^{102} and Ser^{195} are reasonable (Table 3, model T1). Constraining χ_1 and χ_2 to ideality (Table 3, model T2) or rotating the His P2 through 180° (to pair the Nδ1 with Ser^{195} and Nε2 with Asp^{102}) gives less favorable hydrogen-bond distances, angles, and steric contacts.

Model building and experimental data presented above show that the function of the catalytic histidine in subtilisin can be partially replaced by a His side chain from the P2 position of the substrate. We suggest that similar experiments in the trypsin family of proteases may yield similar results. There is no compelling mechanistic reason why a His at the P2 position should be so poised in close proximity to the catalytic histidine in two convergently evolved enzymes. Instead, we speculate that this could represent an evolu-

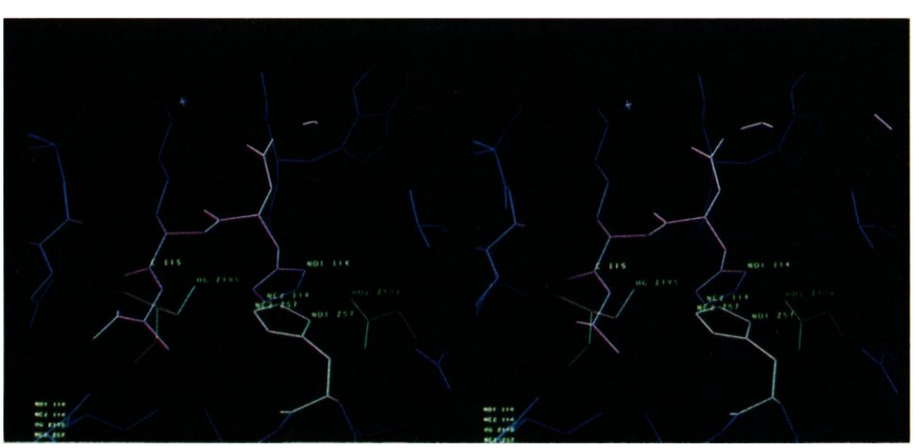

Fig. 4. Stereo view of a complex (*38*) between bovine trypsin (blue) and bovine pancreatic trypsin inhibitor (BPTI) complex (pink) in which the equivalent P2 substrate side chain (Cys^{14} in BPTI) is replaced by His and superimposed upon His^{57} in trypsin. The main chain position of BPTI remained fixed while the dihedral angles for the His side chain (substituted for Cys^{14}) were varied. The catalytic triad of trypsin (Ser^{195}, His^{57}, Asp^{102}) is shown in green and the carbonyl carbon of Lys^{15} at the P1 position in BPTI is labeled (the model corresponds to model T1 in Table 3).

tionary vestige of when serine proteases operated by substrate-assisted catalysis.

Nature provides a possible example of a serine protease in which the catalytic base is donated from another molecule. In the autoproteolytic maturation of the human rhinovirus coat protein VP0 to VP2 and VP4, it has been proposed that an RNA base acts as a proton acceptor-donor by inserting between a catalytic serine and carboxylate (*33*). It is plausible that enzymes may evolve by taking maximal advantage of catalytic groups present in their environment, whether located on a cofactor or on the substrate. Some enzymes may remain dependent upon substrate functional group (or groups), perhaps because the enzymes are too "young," or the substrate functional group (or groups) cannot be adequately mimicked by an amino acid side chain (or a post-translational modification thereof) of the enzyme, or because it serves an important regulatory function. Other enzymes, such as the serine proteases, may eventually dispense with the requirement for a substrate functional group because the latter can be adequately replaced by its own amino acid side chain allowing the enzyme to relax substrate specificity.

REFERENCES AND NOTES

1. M. Smith, *Annu. Rev. Genet.* **19**, 423 (1985); P. Carter, *Biochem. J.* **237**, 1 (1986).
2. D. A. Estell *et al.*, *Science* **233**, 659 (1986).
3. J. A. Wells, D. B. Powers, R. R. Bott, T. P. Graycar, D. A. Estell, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1219 (1987).
4. J. A. Wells, B. C. Cunningham, T. P. Graycar, D. A. Estell, *ibid.*, p. 5167.
5. C. S. Craik *et al.*, *Science* **228**, 291 (1985).
6. A. R. Fersht *et al.*, *Nature (London)* **314**, 235 (1985).
7. J. R. Winther, M. C. Kielland-Brandt, K. Breddam, *Carlsberg Res. Commun.* **50**, 273 (1985).
8. C. Murali and E. M. Creaser, *Prot. Eng.* **1**, 55 (1986).
9. C. N. Cronin, B. A. Malcolm, J. F. Kirsh, *J. Am. Chem. Soc.* **109**, 2222 (1987).
10. R. M. Stroud, M. Krieger, R. E. Koeppe, A. A. Kossiakoff, J. L. Chambers, in *Proteases and Biological Control*, E. Reich, D. B. Rifkin, E. Shaw, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1975), p. 13; J. Kraut, *Annu. Rev. Biochem.* **46**, 331 (1977).
11. Peptide substrate nomenclature may be represented as:

$$NH_2\text{-Pn...P2-P1-}\overset{\overset{O}{\|}}{C}\text{-}\overset{\overset{H}{|}}{N}\text{-P1'-P2'...Pn'-}COOH$$

where the scissile bond is between the P1 and the P1′ substrate residues; see I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.* **27**, 157 (1967).
12. S. D. Power, R. M. Adams, J. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 3096 (1986).
13. The designation His64Ala refers to a mutation where the His residue at position 64 in wild-type subtilisin is mutated to Ala. The designation Ser24Cys:His64Ala refers to a double mutant protein in which the wild-type residues Ser^{24} and His^{64} are converted to Cys and Ala, respectively.
14. The His64Ala mutation was constructed with the use of the oligonucleotide: 5′-CAACAACTCCGCGGGAACTCAC-3′ (asterisks denote mismatches to the wild-type sequence and underlined is a unique Sac II site) and a pSS5 single-stranded plasmid template (B. Cunningham, D. Powers, J. Wells, unpublished) containing the subtilisin gene (*15*) using a previously described method (*16*). The double-mutant Ser24Cys:His64Ala was constructed from the His64Ala and Ser24Cys single mutants by a three-way ligation with the use of the following fragments: 6-kb Eco RI–Bam HI vector fragment from pSS5, 0.5-kb Eco RI–Cla I from Ser24Cys (*17*), and 1.0-kb Cla I–Bam HI from His64Ala. The entire coding sequence of the Ser24Cys:His64Ala mutant was then verified by dideoxy sequencing (*18*). Transformation of *B. subtilis* BG2036 (*19*) with *Escherichia coli–B. subtilis* shuttle plasmids has been described (*17*).
15. J. A. Wells, E. Ferrari, D. J. Henner, D. A. Estell, E. Y. Chen, *Nucleic Acids Res.* **11**, 7911 (1983).
16. P. Carter, H. Bedouelle, G. Winter, *ibid.* **13**, 4331 (1986).
17. J. A. Wells and D. B. Powers, *J. Biol. Chem.* **261**, 6564 (1986).
18. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977).
19. M. Y. Yang, E. Ferrari, D. J. Henner, *J. Bacteriol.* **160**, 15 (1984).
20. P. Carter and J. A. Wells, unpublished results.
21. Separate *B. subtilis* BG2036 cultures, containing the Ser24Cys:His64Ala double-mutant subtilisin plasmid and a functionally silent Ala48Glu mutant subtilisin plasmid (*22*), were grown in 2 × TY media (*23*) containing chloramphenicol (12.5 μg/ml) at 37°C for 18 to 20 hours. Co-cultures were inoculated by diluting Ser24Cys:His64Ala 1 to 100 (v/v) and Ala48Glu 1 to 100,000 (v/v) in 2× TY containing chloramphenicol (12.5 μg/ml) with 10 m*M* $CaCl_2$, and grown at 37°C for 20 to 24 hours with vigorous aeration. Co-cultures (2 liters) were centrifuged (8000*g*, 15 minutes, 4°C) and three volumes of ethanol (−20°C) was added to the supernatant. After centrifugation (8000*g*, 15 minutes, 4°C), the pellet was resuspended in 50 m*M* tris-HCl (*p*H 8.0), 5 m*M* $CaCl_2$, 10 m*M* dithiothreitol (DTT). After centrifugation (40,000*g*, 30 minutes, 4°C), the supernatant was dialyzed against 2 liters of 10 m*M* 2-[*N*-morpholino]ethanesulfonic acid (MES) (*p*H 6.2), 5 m*M* $CaCl_2$, 10 m*M* DTT (S buffer) overnight at 4°C. The dialyzate was passed over a 50-ml DE52 column, and the flow through was placed on a 50-ml CM Trisacryl (LKB) column. Subtilisin was eluted with a 600-ml gradient of S buffer containing NaCl from 0 to 100 m*M*. Pooled subtilisin-containing fractions were dialyzed against 2 liters of deaerated 10 m*M* MES (*p*H 6.2), 5 m*M* $CaCl_2$, 100 m*M* NaCl (T buffer). Samples were loaded onto an activated thiol Sepharose matrix (Pharmacia), washed extensively with T buffer, and eluted with T buffer containing 20 m*M* DTT. The eluate was concentrated with Centricon 10 microconcentrators (Amicon) and dialyzed against four changes of 2 liters of 10 m*M* MES (*p*H 6.2), 5 m*M* $CaCl_2$, 10 m*M* DTT. The concentration of subtilisin was determined from the measured absorbance at 280 nm ($\epsilon^{0.1\%} = 1.17$) (*24*). The Ser24Cys mutant was cultured in the absence of helper and was purified as for Ser24Cys:His64Ala. Purified enzymes were flash frozen in aliquots and stored at −70°C.
22. J. Wells, unpublished data.
23. J. H. Miller, in *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1972).
24. H. Matsubara, C. B. Kaspar, D. M. Brown, E. L. Smith, *J. Biol. Chem.* **240**, 1125 (1965).
25. J. H. Morrissey, *Anal. Biochem.* **117**, 307 (1981); U. K. Laemmli, *Nature (London)* **227**, 680 (1970).
26. D. A. Estell and T. P. Graycar, unpublished results.
27. H. Gutfreund and J. M. Sturtevant, *Biochem. J.* **63**, 656 (1956). For the two-step reaction mechanism of serine proteases,

$$E + S \overset{K_s}{\rightleftharpoons} E \cdot S \overset{k_2}{\underset{Pr_1}{\rightleftharpoons}} E\text{-}Ac \overset{k_3}{\underset{H_2O}{\rightarrow}} E + Pr_2$$

(where E is enzyme, S is substrate, E−Ac is the acyl-enzyme, and Pr_1 and Pr_2 are the products), $K_m = K_s\, k_3/(k_2 + k_3)$. When acylation is rate determining (that is, $k_3 >> k_2$) $K_m \sim K_s$; when deacylation is rate determining (that is, $k_2 >> k_3$) $K_m = K_s\,(k_3/k_2)$.
28. J. A. Wells, B. C. Cunningham, T. P. Graycar, D. A. Estell, *Philos. Trans. R. Soc. London, Ser. A* **317**, 415 (1986). For wild-type enzyme, the ratio of the acylation rate (k_2) to the deacylation rate constant (k_3) is at least 1:33 for the substrate succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide.
29. A. N. Glazer, *J. Biol. Chem.* **242**, 433 (1967); P. G. Thomas, A. J. Russell, A. R. Fersht, *Nature (London)* **318**, 375 (1985).
30. M. Ottesen and I. Svendsen, in *Methods in Enzymology*, C. Perleman, Ed. (Academic Press, New York, 1970), vol. 19, p. 199.
31. M. Philipp and M. L. Bender, *Mol. Cell. Biochem.* **51**, 5 (1983); I. Svendsen, *Carlsberg Res. Commun.* **41**, 237 (1976).
32. M. H. Klapper, *Biochem. Biophys. Res. Commun.* **78**, 1018 (1977).
33. M. G. Rossmann *et al.*, *Nature (London)* **317**, 145 (1985); E. Arnold *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 21 (1987).
34. E. G. Del Mar, C. Largman, J. W. Brodrick, M. C. Goekas, *Anal. Biochem.* **99**, 316 (1979).
35. S. J. Weiner *et al.*, *J. Am. Chem. Soc.* **106**, 765 (1984).
36. T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978).
37. R. Bott and M. Ultsch, unpublished results.
38. J. Deisenhofer and W. Steigemann, *Acta Crystallogr. B.* **31**, 238 (1975). Coordinates for trypsin-trypsin inhibitor complex taken from the Brookhaven Protein Data Bank entry 2PTC deposited by R. Huber and J. Deisenhofer, 9/82.

39. We thank J. Burnier for synthetic peptides, peptide substrates, and helpful discussions, R. Bott for providing crystallographic data prior to publication and help with computer graphics, M. Winkler for helpful discussions, B. Nivens for technical assistance with amino acid analysis, the organic chemistry department at Genentech for oligonucleotide synthesis, W. Anstine for manuscript preparation, and colleagues at Genentech and Genencor for support and encouragement. P.C. is on leave of absence from Gonville and Caius College, Cambridge, U.K.

3 April 1987; accepted 17 June 1987