Toward Tailoring the Specificity of the S_1 Pocket of Subtilisin *B. lentus*: Chemical Modification of Mutant Enzymes as a Strategy for Removing Specificity Limitations†

Grace DeSantis, Xiao Shang, and J. B. Jones*

Department of Chemistry, University of Toronto, 80 Saint George Street, Toronto, Ontario M5S 3H6, Canada

Received April 14, 1999; Revised Manuscript Received July 15, 1999

ABSTRACT: In both protein chemistry studies and organic synthesis applications, it is desirable to have available a toolbox of inexpensive proteases with high selectivity and diverse substrate preferences. Toward this goal, we have generated a series of chemically modified mutant enzymes (CMMs) of subtilisin *B. lentus* (SBL) possessing expanded S_1 pocket specificity. Wild-type SBL shows a marked preference for substrates with large hydrophobic P_1 residues, such as the large Phe P_1 residue of the standard suc-AAPF-pNA substrate. To confer more universal P_1 specificity on S_1, a strategy of chemical modification in combination with site-directed mutagenesis was applied. For example, WT-SBL does not readily accept small uncharged P_1 residues such as the $-CH_3$ side chain of alanine. Accordingly, with a view to creating a S_1 pocket that would be of reduced volume providing a better fit for small P_1 side chains, a large cyclohexyl group was introduced by the CMM approach at position S166C with the aim of partially filling up the S_1 pocket. The S166C-S-CH_2-*c*-C_6H_{11} CMM thus created showed a 2-fold improvement in k_{cat}/K_M with the suc-AAPA-pNA substrate and a 51-fold improvement in suc-AAPA-pNA/suc-AAPF-pNA selectivity relative to WT-SBL. Furthermore, WT-SBL does not readily accept positively or negatively charged P_1 residues. Therefore, to improve SBL's specificity toward positively and negatively charged P_1 residues, we applied the CMM methodology to introduce complementary negatively and positively charged groups, respectively, at position S166C in S_1. A series of mono-, di-, and trinegatively charged CMMs were generated and all showed improved k_{cat}/K_Ms with the positively charged P_1 residue containing substrate, suc-AAPR-pNA. Furthermore, virtually arithmetic improvements in k_{cat}/K_M were exhibited with increasing number of negative charges on the S166C-R side chain. These increases culminated in a 9-fold improvement in k_{cat}/K_M for the suc-AAPR-pNA substrate and a 61-fold improvement in suc-AAPR-pNA/suc-AAPF-pNA selectivity compared to WT-SBL for the trinegatively charged S166C-S-$CH_2CH_2C(COO^-)_3$ CMM. Conversely, the positively charged S166C-S-$CH_2CH_2NH_3^+$ CMM generated showed a 19-fold improvement in k_{cat}/K_M for the suc-AAPE-pNA substrate and a 54-fold improvement in suc-AAPE-pNA/suc-AAPF-pNA selectivity relative to WT-SBL.

For both protein chemistry (*1*−*3*) and organic synthesis applications (*4*−*6*), it is desirable to have available a diverse toolbox of inexpensive proteases with high selectivity and diverse substrate preferences. To date, the most extensively exploited class of enzymes in organic synthesis applications has been the hydrolases. Among these, the serine proteases have received considerable attention due, in part, to their often exquisite stereo-, regio-, and chemoselectivities (*4*−*7*). While over 3000 enzymes have now been reported, of which many are proteases, significantly fewer of the latter are available inexpensively from commercial sources (*5, 8*). Furthermore, since WT[1] enzymes will never accept all substrate structures of synthetic interest, it is attractive to contemplate the tailoring of a readily available protease in order to expand its substrate specificity in a controlled manner with the ultimate goal of creating any desired specificity at will.

In this regard, the goal of specificity alteration of enzymes has already been targeted by several different approaches. For example, site-directed mutagenesis (*9*) and random mutagenesis (*10*) have been employed to tailor enzyme specificity and have permitted some insights into the electrostatic (*11*−*16*), steric (*17*−*22*), and hydrophobic (*20, 23*−*24*) factors which govern enzyme−substrate interactions. In particular, the directed evolution approach to altering enzyme specificity is now beginning to make significant advances toward these goals (*10*). However, the structural variations within these approaches are limited to the 20

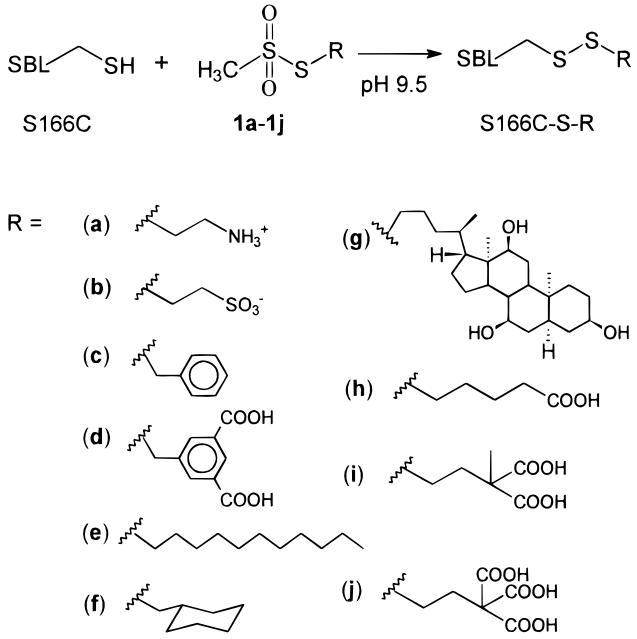
† We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) and Genencor International, Inc. for generous financial support and NSERC for scholarship support (to G.D.).

* To whom correspondence should be addressed. E-mail: jbjones@alchemy.chem.utoronto.ca. Phone: (416) 978-3589. Fax: (416) 978-1553.

[1] Abbreviations: Ches, 2-(cyclohexylamino)ethanesulfonic acid; CMM, chemically modified mutant enzyme; DMSO, dimethyl sulfoxide; ES-MS, electrospray mass spectrometry; EtOH, ethanol; FPLC, fast-performance liquid chromatography; Mes, 4-morpholineethanesulfonic acid; MTS, methanethiosulfonate; PAGE, polyacrylamide gel electrophoresis; PEG, poly(ethylene glycol); SBL, subtilisin Bacillus lentus; SDM, site-directed mutagenesis; Suc-AAPF-pNA succinyl-alanyl-alanyl-prolyl-phenylalanyl-para-nitroanilide; Tris, tri(hydroxyl methyl)aminomethane; TFA, trifluoroacetic acid; WT, wild-type.

10.1021/bi990861o CCC: $18.00
Published on Web 09/14/1999

Scheme 1



natural amino acids. Consequently, biosynthetic methods have recently been developed to introduce unnatural amino acids into proteins (*25*−*27*). Unnatural functionalities have also been incorporated by chemical modification techniques (*28*−*34*). Since the unnatural amino acid mutagenesis approach is not yet amenable to large-scale preparations and chemical modification alone is insufficiently specific, we have begun to exploit a strategy of applying a combination of site-directed mutagenesis and chemical modification to modify enzyme specificity (*35*−*40*). This approach is illustrated in Scheme 1 and entails the introduction of a unique cysteine residue at a selected position, followed by its chemical modification with methanethiosulfonate (*41*−*43*) reagents (MTS, **1a**−**j**) to generate chemically modified mutant enzymes (CMMs). The combination of site-directed mutagenesis and chemical modification has previously been recognized as a powerful tool for the creation of new active-site environments (*19*, *44*), in mechanistic studies (*45*−*46*) for the investigation of protein packing (*47*), and for cofactor incorporation (*28*). This approach has also been applied to detailed studies of ion-channel properties (*48*−*50*) for site-directed introduction of spin labels (*51*, *52*), to probe receptor binding (*53*), and in investigations of membrane spanning proteins (*54*, *55*).

The subtilisin from *Bacillus lentus* (SBL, EC 3.4.21.14) is well suited as an exploratory vehicle for evaluating the potential of this combined site-directed mutagenesis chemical modification approach since it is a well-characterized enzyme and is of synthetic (*56*, *57*) as well as industrial (*58*) interest. Furthermore, SBL's high-resolution crystal structure has been solved (*59*, *60*), and it has been cloned, overexpressed, and purified (*61*), and its kinetic behavior well characterized (*62*−*65*). In addition, and importantly, wild-type (WT) SBL contains no natural cysteine residues, and methanethiosulfonate reagents therefore react *only* with the introduced cysteine residue. The validity of the CMM approach for altering the stability (*66*), specificity (*24*), kinetic properties (*35*−*37*, *39*, *40*), and pH profiles (*38*) of subtilisins has been recognized.

WT-SBL has a marked preference for substrates with large hydrophobic uncharged P_1 residues. In this study, we explore tailoring of the S_1 pocket of SBL to also accept small hydrophobic, positively charged and negatively charged P_1 residues. To achieve this broadened P_1 tolerance, a simplistic strategy of steric and electrostatic complementarity was applied (*67*). Employing the crystal structure of SBL as our guide (*59*), the Ser166 residue, which is located at the bottom of the S_1 pocket and whose side chain points inward toward the pocket, was chosen for mutagenesis to cysteine and subsequent chemical modification. First, to expand SBL's specificity toward small uncharged P_1 residues, such as the small P_1 Ala residue of the suc-AAPA-pNA substrate, we introduced large moieties at position 166 in S_1, such as benzyl (**-c**), decyl (**-e**), cyclohexyl (**-f**), and steroidyl (**-g**) groups with a view to reducing the volume of S_1 and inducing a better fit of small P_1 groups, thereby conferring elastase-like (*68*) substrate specificity on SBL. Then, to expand SBL's specificity toward positively charged P_1 residues, such as the P_1 Arg residue of the suc-AAPR-pNA substrate, we introduced negatively charged groups at position S166C in S_1, such as the ethylsulfonato (**-b**) moiety, and the dicarboxylic aromatic (**-d**) and aliphatic mono- (**h**), di- (**i**), and tri (**-j**)-aliphatic groups, to elicit complementary electrostatic attractions with a view to making SBL trypsin-like in its specificity (*69*). Conversely, to expand SBL's specificity toward negatively charged P_1 residues, such as the negatively charged P_1 Glu residue of the suc-AAPE-pNA substrate, we introduced the positively charged ethylamino (**-a**) group at position S166C in S_1.

RESULTS

The preparations of the requisite MTS reagents **1c** (*36*), **1e**,**f** (*36*), **1d** (*36*, *70*), and **1i**,**j** (*36*, *70*) are as reported previously, and the steroidyl MTS reagent **1g** was prepared from cholic acid by the same methodology.

Each of the CMMs obtained was characterized in order to establish its purity and integrity. Titration of the CMMs with Ellman's reagent showed a residual thiol content of less than 2% in all cases, demonstrating that the MTS reactions were virtually quantitative. Mass analyses of the CMMs by electrospray mass spectrometry were consistent (± 6 Da) with the calculated masses. The purities of the modified enzymes were assessed by native-PAGE, and in all cases, only one band was visible. Furthermore, as expected relative to WT, the negatively charged CMMs S166C-S-**b**, **-d**, and **-i** to **-j** displayed retarded mobility in the direction of the cathode, while the positively charged S166C-S-**a** CMM displayed greater mobility. That modification of cysteine is wholly responsible for altered activity was established by the absence of reaction of WT-SBL with the MTS reagents. Also, the modifications are fully reversible by treatment of each of the CMMs with β-mercaptoethanol, further verifying that chemical modification at cysteine was solely responsible for the observed changes in activity. The total amount of active enzyme was determined by titration with phenylmethanesulfonyl fluoride (*71*).

Initially, three CMMs S166C-S-**a**, **-b**, and **-c**, with a positive, a negative, and with a large hydrophobic side chain, respectively, were subjected to a k_{cat}/K_M screen with each of the test substrates, suc-AAP-F/A/R/E-pNA, in order to

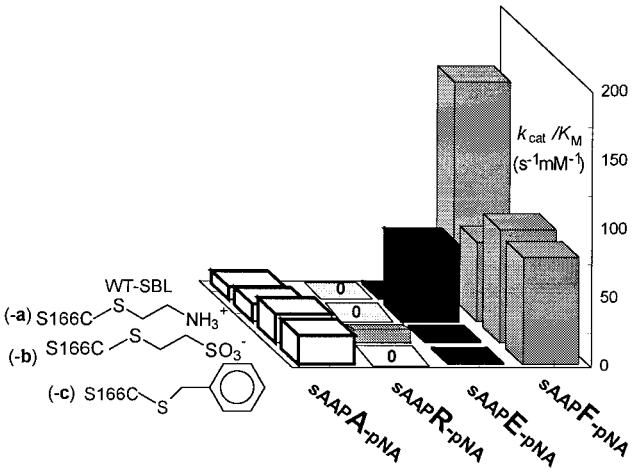


FIGURE 1: k_{cat}/K_M screen of WT-SBL and S166C-S-**a** to -**c** CMMs with each of the standard suc-AAPF-pNA substrate and with the suc-AAP-A/R/E-pNA substrates.

identify any induced complementary electrostatic or improved hydrophobic interactions (Figure 1). While, as expected, k_{cat}/K_Ms with the standard suc-AAPF-pNA were lowered, the k_{cat}/K_Ms of the CMMs whose S_1 sites were tailored toward the Ala, Arg, and Glu P_1 residues, improved with the appropriate substrate. This is illustrated in Figure 1 in the higher activity of S166C-S-**c** with suc-AAPA-pNA, of S166C-S-**b** with suc-AAPR-pNA, and of S166C-S-**a** with suc-AAPE-pNA, all relative to WT.

Following the validation of the general design strategy from this initial screen, more complete kinetic analyses were undertaken. The substrate specificity of each of the CMMs was evaluated kinetically with the standard large hydrophobic P_1 residue containing substrate suc-AAPF-pNA. In addition, the S166C CMMs modified with the large hydrophobic MTS reagents **1c**,**e**−**g**, were evaluated with the small hydrophobic P_1 residue containing substrate, suc-AAPA-pNA. The S166C CMMs modified with the negatively charged MTS reagents **1b**,**d**,**h**−**j** were evaluated with the positively charged P_1 residue containing substrate, suc-AAPR-pNA. The S166C CMM modified with the positively charged MTS reagent **1a** was evaluated with the negatively charged P_1 residue containing substrate, suc-AAPE-pNA. The results are summarized in Table 1.

DISCUSSION

The significant substrate preference of WT-SBL for large hydrophobic P_1 residues is apparent from its preference for the Phe P_1 residue of the standard suc-AAPF-pNA substrate, by a factor of 9500-fold over the small P_1 residue of suc-AAPA-pNA, by a factor of 24-fold compared to the positively charged P_1 residue of suc-AAPR-pNA and by a factor of 522-fold compared to the negatively charged P_1 residue of suc-AAPE-pNA (Table 1, entries 1, 12, 17, and 23). These kinetic differences are due to changes in both binding, as reflected by K_M, and in turnover number, k_{cat}. Moreover, and predictably, the WT enzyme is by far the best catalyst with suc-AAPF-pNA, and its conversions to any of the CMMs were deleterious with respect to this substrate and resulted in k_{cat}/K_M decreases of up to 34-fold (Table 1, entries 2−11).

To improve the substrate specificity of SBL toward small hydrophobic P_1 residues such as Ala, the simplistic approach of filling up the S_1 binding cleft was addressed by preparing the S166C-S-$CH_2C_6H_5$ (-**c**), S166C-S-$CH_2(CH_2)_8CH_3$ (-**e**), S166C-S-$CH_2C_6H_{11}$ (-**f**), and S166C-S-steroidyl (-**g**) CMMs. This design strategy attempted to mimic the function of the bulky S_1-pocket side chains of α-lytic protease (*72*−*74*) and of elastase (*68*), which are responsible for their substantial preference for the small P_1 group containing suc-AAPA-pNA substrate over the large P_1-group containing suc-AAPF-pNA substrate (*68*, *72*−*74*). These CMMs (S166C-S-**c**,-**e**, -**f**, -**g**. Table 1, entries 12−16) were then evaluated with the suc-AAPA-pNA substrate. All revealed slightly improved binding compared to WT, with the greatest improvement in K_M being 2-fold for the S166C-S-$CH_2C_6H_5$ (-**c**) CMM. However, of these four CMMs, only S166C-S-$CH_2C_6H_{11}$ (-**f**) showed both an improved k_{cat} and an improved k_{cat}/K_M. While this design strategy yielded only one CMM with an increased preference for the small Ala P_1 residue, all of these modifications effectively excluded the larger Phe P_1 residue preferred by WT-SBL (Table 1, entries 4 and 6−8). Overall, the selectivities with respect to k_{cat}/K_M for the suc-AAPA-pNA substrate compared to the suc-AAPF-pNA substrate were improved by 11-fold for S166C-S-$CH_2C_6H_5$ (-**c**), 1.1-fold for S166C-S-$CH_2(CH_2)_8CH_3$ (-**e**), 51-fold for S166C-S-$CH_2C_6H_{11}$ (-**f**), and 3.2-fold for S166C-S-steroidyl (-**g**), all compared to WT. These differences in P_1 Ala selectivity may be a reflection of the orientation of the R side chain of the CMM, with the side chains of S166C-S-$CH_2C_6H_5$ (-**c**) and −$CH_2C_6H_{11}$ (-**f**), behaving as though directed into the pocket and favoring P_1 = Ala, whereas the side chains of S166C-S-$CH_2(CH_2)_8CH_3$ (-**e**) and S166C-S-steroidyl (-**g**) behave as though directed outward, thus not significantly altering the shape of the S_1 pocket.

The above improvements in P_1 Ala acceptance, although modest, are encouraging and demonstrate the effectiveness of the adopted strategy (*20*−*22*, *75*) Tailoring the steric complementarity between enzymes and substrates has already been found to be challenging, and comparison of these CMM results with previous literature studies targeting the same goal are both interesting and intriguing. For example, the G166I mutation of subtilisin BPN′ effected an almost 1000-fold decrease in k_{cat}/K_M with the P_1=Phe suc-AAPF-pNA substrate. However, this same G166I mutant, which was by far the most effective mutation, elicited a 10-fold improved k_{cat}/K_M compared to WT with the P_1=Ala suc-AAPA-pNA substrate (*20*). Notably, the G166I subtilisin mutant is more selective for Ala over Phe than our most selective CMM. However, in both cases, the increases in selectivity are due mainly to decrease in k_{cat}/K_M for suc-AAPF-pNA rather than increases in k_{cat}/K_M for suc-AAPA-pNA. The results for both CMM and SDM approaches agree that decreasing the selectivity of an enzyme for a large hydrophobic residue containing substrate can be accomplished in a relatively facile manner by the introduction of large amino acid in the enzyme pocket, but that increasing the selectivity of an enzyme for a small hydrophobic residue containing substrate is much more difficult. Similarly, the G127A mutant of subtilisin YaB, whose specificity was already elastase-like, effected a 10-fold improvement in k_{cat}/K_M with the suc-AAPA-pNA substrate (*75*). However, the G127V mutant of subtilisin E induced a decrease in k_{cat}/K_M with the suc-AAPA-pNA substrate, identifying an inconsistency in the SDM strategy (*21*). Thus, the CMM approach offers a complementary

Table 1: Kinetic Evaluation[a] of Altered S_1 Pocket Specificity

entry	enzyme	substrate	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}\,mM^{-1}$)
1	WT	suc-AAPF-pNA	0.73 ± 0.08	153 ± 4	209 ± 15
2	S166C-S-**a**	suc-AAPF-pNA	0.68 ± 0.04	50 ± 1	74 ± 5
3	S166C-S-**b**	suc-AAPF-pNA	1.34 ± 0.08	25.0 ± 0.7	19 ± 1
4	S166C-S-**c**	suc-AAPF-pNA	1.17 ± 0.06	23.1 ± 0.5	20 ± 1
5	S166C-S-**d**	suc-AAPF-pNA	1.6 ± 0.2	47 ± 3	29 ± 4
6	S166C-S-**e**	suc-AAPF-pNA	1.09 ± 0.07	82 ± 2	75 ± 5
7	S166C-S-**f**	suc-AAPF-pNA	0.70 ± 0.05	4.8 ± 0.1	6.90 ± 0.05
8	S166C-S-**g**	suc-AAPF-pNA	0.74 ± 0.07	29 ± 1	41 ± 4
9	S166C-S-**h**[b]	suc-AAPF-pNA	1.52 ± 0.06	48 ± 1	31 ± 1
10	S166C-S-**i**[b]	suc-AAPF-pNA	2.26 ± 0.10	67 ± 2	30 ± 2
11	S166C-S-**j**[b]	suc-AAPF-pNA	2.46 ± 0.11	76 ± 2	31 ± 2
12	WT	suc-AAPA-pNA	2.0 ± 0.1	17.7 ± 0.3	8.8 ± 0.4
13	S166C-S-**c**	suc-AAPA-pNA	0.8 ± 0.1	6.8 ± 0.3	9 ± 1
14	S166C-S-**e**	suc-AAPA-pNA	1.90 ± 0.03	6.8 ± 0.4	3.6 ± 0.6
15	S166C-S-**f**	suc-AAPA-pNA	1.90 ± 0.07	28.2 ± 0.4	14.8 ± 0.6
16	S166C-S-**g**	suc-AAPA-pNA	1.74 ± 0.04	9.65 ± 0.07	5.54 ± 0.3
17	WT	suc-AAPR-pNA	7.2 ± 0.7	0.16 ± 0.01	0.022 ± 0.002
18	S166C-S-**b**	suc-AAPR-pNA	3.4 ± 0.3	0.17 ± 0.01	0.050 ± 0.005
19	S166C-S-**d**	suc-AAPR-pNA	5.5 ± 1.1	0.68 ± 0.08	0.12 ± 0.03
20	S166C-S-**h**	suc-AAPR-pNA	8.2 ± 0.9	0.35 ± 0.02	0.041 ± 0.005
21	S166C-S-**i**	suc-AAPR-pNA	5.3 ± 0.5	0.43 ± 0.02	0.080 ± 0.008
22	S166C-S-**j**	suc-AAPR-pNA	5.2 ± 0.6	1.06 ± 0.07	0.20 ± 0.03
23	WT	suc-AAPE-pNA	4.4 ± 0.4	1.75 ± 0.08	0.40 ± 0.04
24	S166C-S-**a**	suc-AAPE-pNA	1.9 ± 0.1	14.5 ± 0.3	7.6 ± 0.4

[a] Michaelis−Menten constants were measured by the initial rates method in, pH 8.6, Tris-HCl buffer at 25 °C with suc-AAPF-pNA as the substrate. [b] Taken from ref *70*.

alternative to conventional site-directed mutagenesis toward the goal of tailoring the steric complementarity between enzymes and substrates (*22*).

Improving the substrate specificity of SBL toward positively charged P_1 residues such as Arg was based on mimicking the common motif in trypsin-like enzymes (*69*) of high negative charge density of acidic residues that favor binding of positively charged substrate structures (*69*, *76*−*78*). This goal was addressed by S166C-S-$CH_2CH_2SO_3^-$ (-**b**), S166C-S-$CH_2(CH_2)_2CH_2COO^-$ (-**h**), S166C-S-$CH_2C_6H_4$-3,5-$(COO^-)_2$ (-**d**), S166C-S-$CH_2CH_2C(CH_3)(COO^-)_2$ (-**i**), and S166C-S-$CH_2CH_2C(COO^-)_3$ (-**j**), a series of CMMs which provide an S_1 pocket which is potentially mono-, di-, and trinegatively charged. Evaluation of each of these CMMs with the suc-AAPR-pNA substrate revealed K_Ms that were up to 2-fold improved compared to WT (Table 1, entries 17−22). The general success of this approach is evident since all of the CMMs with a negatively charged −R side chain showed improved activity compared to WT with an up to 7-fold improved k_{cat} and an up to 9-fold improved k_{cat}/K_M with the suc-AAPR-pNA substrate (Table 1, entries 17−22). Overall, the selectivities, with respect to k_{cat}/K_M, for the suc-AAPR-pNA substrate compared to the suc-AAPF-pNA substrate were improved 25-fold for S166C-S-$CH_2CH_2SO_3^-$ (-**b**), 13-fold for S166C-S-$CH_2(CH_2)_2CH_2COO^-$ (-**h**), 39-fold for S166C-S-$CH_2C_6H_4$-3,5-$(COO^-)_2$ (-**d**), 25-fold for S166C-S-$CH_2CH_2C(CH_3)(COO^-)_2$ (-**i**), and 61-fold for S166C-S-$CH_2CH_2C(COO^-)_3$ (-**j**) relative to WT.

The strategy of introducing charge complementarity to induce trypsin-like P_1 specificity in subtilisins has previously been explored using site-directed mutagenesis (SDM) (*12*, *16*, *79*). Interestingly, however, the G166D and G166E mutants of subtilisin BPN′ caused decreases in k_{cat}/K_M with the suc-AAPK-pNA substrate rather than the anticipated increases (*12*). Furthermore, while the G166D subtilisin BPN′ mutant was reported to exhibit a 18-fold improvement in Arg/Phe P_1 selectivity, this was accompanied by a 2.5-fold decrease in k_{cat}/K_M with the suc-AAPR-pNA substrate compared to WT (*16*). Thus, compared to WT, both with respect to improved k_{cat}/K_M with the suc-AAPR-pNA substrate and improved Arg/Phe P_1 selectivity, the CMMs reported herein are more successful.

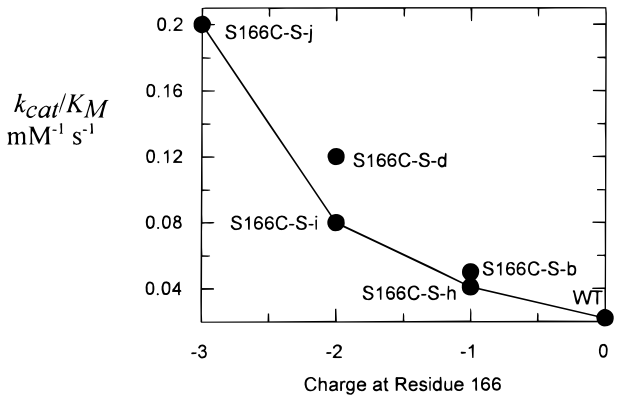


FIGURE 2: Plot of side-chain negative charge on R at position 166 versus k_{cat}/K_M with suc-AAPR-pNA. The points on the line are for the aliphatic series of carboxylated −R groups and the WT, Ser166.

Furthermore, these improvements correlate directly with the number of introduced negative charges, such that each additional negative charge introduced at position 166 causes an approximate doubling in k_{cat}/K_M with the complementary charged suc-AAPR-pNA substrate (Figure 2). This arithmetic improvement is interesting since it contrasts with the previous observation that while remote noninteracting charged mutations are additive, multiple interacting adjacent charged mutations often are not additive. Rather, in most cases, the empirically observed effects are lower than predicted from their individual sums (*80*). This phenomenon is particularly problematic for charged mutations due to the long-range effects of electrostatic interactions. For example, the sum of the transition-state stabilization energy, $\Delta\Delta G_T^{\ddagger}$, for the two

single positively charged subtilisin mutations (D99K and E156K) over estimates the empirically observed effect of the double mutant, when assayed with an Arg P_1 containing substrate (*80*). In contrast, in all cases the empirically determined $\Delta\Delta G^{\ddagger}_T$ values for the aliphatic carboxylate series of mono- di- and trinegatively charged side chains, of the S166C-S-**h**, -**i**, and -**j** CMMs exhibit an additive effect of additional charge (*81*). Thus, the CMM approach offers a convenient method to circumvent the problem of the attenuation of the augmenting effect of the introduction of additional charges by SDM by permitting the introduction of a larger local charge density.

Conversely, the adopted design strategy of introducing a complementary positive charge in the S_1 binding cleft by the CMM approach to improve P_1 = Glu selectivity was based on mimicking the specificity determinants of the serine proteases Pronase (*82*−*83*) and granzyme B (*84*−*86*), which exhibit a substrate preference for negatively charged P_1 residues, and whose S_1 pockets are lined with positively charged residues. The success of the current approach is apparent from the remarkable 19-fold increase in k_{cat}/K_M, with the suc-AAPE-pNA substrate displayed by S166C-S-CH_2-$CH_2NH_3^+$ (-**a**). This enhancement is due to a combination of better binding, evident from the 2-fold lower K_M and 8-fold higher k_{cat} (Table 1, entries 23 and 24). The induction of electrostatic complementarity was most unequivocally demonstrated by the 54-fold improvement in suc-AAPE-pNA to suc-AAPF-pNA substrate selectivity, with respect to k_{cat}/K_M, for S166C-S-$CH_2CH_2NH_3^+$ (-**a**) compared to WT. Previously, the individual G166R and G166K subtilisin BPN′ mutations elicited 23- and 340-fold improvements in k_{cat}/K_M for the suc-AAPE-pNA substrate (*12*, *14*). However, it must be noted that the E156Q−G166K double mutant was much more receptive to Glu P_1 and exhibited a 1900-fold improvement compared to WT (*12*). Interestingly, both the G166R and G166K mutants displayed even higher k_{cat}/K_Ms with the hydrophobic P_1 residue containing substrates suc-AAPN-pNA and suc-AAPM-pNA and even with the positively charged P_1 residue containing substrate suc-AAPK-pNA(*12*). While S166C-S-$CH_2CH_2NH_3^+$ (-**a**) still exhibits a 10-fold preference for suc-AAPF-pNA compared to suc-AAPE-pNA (Table 1, entry 2), the substrate screen (Figure 1) shows that suc-AAPA-pNA and suc-AAPR-pNA are poorer substrates.

Since the S166C-S-$CH_2CH_2NH_3^+$ (-**a**) and suc-AAPE-pNA CMM-substrate pair exhibited the greatest k_{cat}/K_M improvement relative to WT, at 19-fold (Table 1, entry 24), more detailed insights into the molecular basis of their interaction was sought using molecular modeling. Using the modeling approach reported previously (*37*), the product inhibitor, AAPE bound to WT-SBL and to the S166C-S-$CH_2CH_2NH_3^+$ (-**a**) CMM was minimized. Molecular modeling revealed that the minimized binding conformations of AAPE to both the WT and S166C-S-$CH_2CH_2NH_3^+$ enzymes are quite similar despite the 19-fold difference in k_{cat}/K_M. However, the ammonium moiety of the S166C-S-CH_2CH_2-NH_3^+ CMM side chain is oriented toward the carboxylate of the glutamic acid P_1 residue, and although it is not quite within salt-bridge distance (N^+ -to- ^-OOC, 4.76 Å), this additional favorable Coulombic interaction between the ammonium side chain of S166C-S-$CH_2CH_2NH_3^+$ and the carboxylate of the glutamic acid P_1 residue is deemed responsible for the observed 19-fold improvement in k_{cat}/K_M for this CMM-substrate pair, compared to WT.

CONCLUSION

While much remains to be done, the overall data clearly provides an encouraging initial validation of the practical effectiveness of the CMM strategy for generating complementary electrostatic and steric enzyme−substrate interactions. For each of the Ala, Arg, and Glu P_1 residues, at least one, and up to five, of the designed CMMs exhibit improved k_{cat}/K_Ms compared to WT. The CMM approach is complementary to the SDM approach and also offers the additional opportunity for the introduction of multiply charged side chains generating high charge densities at single active-site locations. The beneficial effects of the introduction of a localized high charge density was demonstrated by the arithmetic increases in k_{cat}/K_M, with the suc-AAPR-pNA substrate, induced by incremental increases in the negative charge of the S_1 pocket. Further studies toward refining this general strategy are in hand and will be reported in due course.

EXPERIMENTAL SECTION

Sulfonatoethyl methanethiosulfonate (**1a**) and ethylammonium methanethiosulfonate (**1b**) were purchased from Toronto Research Chemicals (2 Brisbane Rd., Toronto, ON, Canada). Reagents **1c**−**f** (*36*) and **1h**−**j** (*70*) were prepared as previously described. ES-MS data were acquired using a PE SCIEX API III Biomolecular mass spectrometer. The tetrapeptide substrates suc-AAPF/A/R/E-pNA were purchased from Bachem Bioscience Inc. (Torrance, CA). All buffer solutions were made up in deionized water.

Site-Specific Chemical Modification. To 25 mg of a S166C mutant, purified as previously described (*37*, *61*) and stored flash frozen in CHES buffer (2.5 mL; 70 mM Ches, 5 mM Mes, and 2 mM $CaCl_2$, pH 9.5) at 20 °C, was added one of the methanethiosulfonate reagents (**1a**−**g**) (100 μL of a 0.2 M solution), in a PEG (10 000)-coated polypropylene test tube, and the mixture agitated in an end-over-end rotator. Blank reactions containing 100 μL of solvent instead of the reagent solution were run in parallel. Each of the modification reactions was monitored spectrophotometrically (ϵ_{410} = 8800 M^{-1} cm^{-1}) (*79*) on a Perkin-Elmer Lambda 2 spectrophotometer, by specific activity measurements. After the reaction was quenched by dilution in MES buffer (5 mM Mes and 2 mM $CaCl_2$, pH 6.5) at 0 °C, the specific activity of the CMM (10 μL), was determined in buffer containing 0.1 M Tris, pH 8.6, 0.005% Tween 80, and 1% DMSO, with the suc-AAPF-pNA substrate (1 mg/mL) at 25 °C. The reaction was terminated when the addition of a further 100 μL of methanethiosulfonate solution effected no further change in specific activity, generally in 30 min to 3 h. The reaction solution was purified on a disposable desalting column (Pharmacia Biotech PD-10, Sephadex G-25 M) preequilibrated with Mes buffer (5 mM Mes and 2 mM $CaCl_2$, pH 6.5) then dialyzed against 20 mM Mes and 1 mM $CaCl_2$, pH 5.8 (3 × 1 L) at 4 °C and aliquoted into 0.5−1.5 mL volumes, flash frozen in liquid nitrogen, and then stored at −20 °C. Modified enzymes were analyzed by nondenaturing gradient (8 to 25%) gels at pH 4.2, run toward the cathode on the Pharmacia Phast-System, and appeared as one single band.

Electrospray Mass Spectrometry. Prior to ES-MS analysis, CMMs were purified by FPLC (Bio-Rad, Biologic System) on a Source 15 RPC matrix (17-0727-20 from Pharmacia) with 5% acetonitrile, 0.01% TFA as the running buffer and eluted with 80% acetonitrile, 0.01% TFA in a one step gradient. (Mass) WT: calcd, 26 698; found, 26 694 (*36*). S166C-S-**a**: calcd, 26 714; found, 26 708 (*36*). S166C-S-**b**: calcd, 26 853; found, 26 851 (*36*). S166C-S-**c**: calcd, 26 836; found, 26 832 (*36*). S166C-S-**d**: calcd, 26 924; found, 26 928 (*70*). S166C-S-**e**: calcd, 26 886; found, 26 890 (*36*). S166C-S-**f**: calcd, 26 842; found, 26 844 (*36*). S166C-S-**g**: calcd, 27 128; found, 27 123 (*36*). S166C-S-**h**: calcd, 26 846; found, 28 646 (*70*). S166C-S-**i**: calcd, 26 890; found, 26 894 (*70*). S166C-S-**j**: calcd, 26 934; found, 26 939 (*70*).

Regeneration of Unmodified Enzyme by Treatment with β-Mercaptoethanol. To a solution of CMM (2.0 mg) in 250 μL of Ches buffer (70 mM Ches, 5 mM Mes, and 2 mM $CaCl_2$, pH 9.5) was added 10 μL of a solution of β-mercaptoethanol (1 M in 95% EtOH). The reaction was monitored by specific activity measurements.

Fee Thiol Titration. The free thiol content of S166C CMMs was determined spectrophotometrically by titration with Ellman's reagent (ϵ_{412} = 13 600 M^{-1} cm^{-1}) (*87*) in phosphate buffer 0.25 M, pH 8.0.

Active-Site Titrations. The active enzyme concentration was determined as previously described (*71*) by monitoring fluoride release upon enzyme reaction with phenylmethanesulfonyl fluoride (Aldrich Chemical Co. Inc.) as measured by a fluoride ion sensitive electrode (Orion Research 96-09). The active enzyme concentration determined in this way was used to calculate kinetic parameters for each CMM.

Kinetic Measurements. Michaelis−Menten constants were measured at 25 °C by curve fitting (GraFit 3.03) of the initial rate data determined at eight concentrations (0.125−8.0 mM) of the suc-AAPX-pNA substrate in, pH 8.6, 0.1 M Tris-HCl buffer containing 0.005% Tween 80 and 1% DMSO (ϵ_{410} = 8800 M^{-1} cm^{-1}).

Molecular Modeling. The X-ray structure of subtilisin *Bacillus lentus* (*59*) was used as the starting point for calculations on the wild-type and chemically modified mutant enzymes. The enzyme setup was performed with Insight II (*88*). To create initial coordinates for the minimization, hydrogens were added at the pH used for kinetic measurements. This protonated all Lys and Arg residues and the N-terminus and deprotonated all Glu and Asp residues and the C-terminus. In addition, the active-site His64 was protonated. The model system with the Ala-Ala-Pro-Phe (from crystal structure) (*59*) product inhibitor bound in the S_1-S_4 pocket was solvated with a 5 Å layer of water molecules giving a total number of water molecules of 1143 in this system. The overall charge of the enzyme−inhibitor complex resulting from this setup was +4 for the WT enzyme. Energy simulations were performed with the Discover program (*89*), on a Silicon Graphics Iris Indigo computer, using the consistent valence force field function (CVFF). A nonbonded cutoff distance of 18 Å with a switching distance of 2 Å was employed. The nonbonded pair list was updated every 20 cycles, and a dielectric constant of 1 was used in all calculations. The WT enzyme was minimized in stages, with initially only the water molecules being allowed to move, then the water molecules and the amino acid side chains, and then the entire enzyme. The mutated and chemically modified enzymes were generated using the Builder module of Insight. Then the amino acid side chains within a 10 Å radius of the α-carbon of the mutated residue were minimized while all other residues were constrained, then all of the atoms within a 10 Å shell were minimized, followed by minimization of the whole system. To examine the effect of a different P_1 residue (Glu), the Phe to Glu mutation of the product inhibitor was constructed using insightII, and then this structure was minimized as above.

ACKNOWLEDGMENT

We are indebted to Genencor International Inc., for supplying us with the WT, and the N62C mutant of SBL, to Drs. Rick Bott, Thomas Graycar, Colin Mitchinson, Christian Paech, and Ben G. Davis for invaluable discussions and to BGD for conducting an initial k_{cat}/K_M quick screen of the S166C-S-**h** to -**j** CMMs.

SUPPORTING INFORMATION AVAILABLE

Preparation of steroidyl methane thiosulfonate reagent, **1g**. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Nilsson, B., Forsberg, G., Moks, T., Hartmanis, M., and Uhlén, M. (1992) *Curr. Opin. Struct. Biol. 2*, 569−575.
2. LaVallie, E. R., and McCoy, J. M. (1995) *Curr. Opin. Biotechnol. 6*, 501−506.
3. Uhlen, M., and Moks, T. (1990) *Methods Enzymol. 185*, 129−143.
4. Sears, P., and Wong, C.-H. (1996) *Biotechnol. Prog. 12*, 423−433.
5. Faber, K. (1997) *Biotransformations in Organic Synthesis*, Springer-Verlag, Heidelberg.
6. Roberts, S. M. (1993) *Preparative Biotransformations*, Wiley, New York.
7. Moree, W. J., Sears, P., Kawashiro, K., Witte, K., and Wong, C.-H. (1997) *J. Am. Chem. Soc. 119*, 3942−3947.
8. Jones, J. B. (1986) *Tetrahedron 42*, 3351−3403.
9. Perona, J. J., and Craik, C. S. (1995) *Protein Sci. 4*, 337−360.
10. Arnold, F. H., and Volker, A. A. (1999) *Curr. Opin. Chem. Biol. 3*, 54−59.
11. Wells, J. A., Cunningham, B. C., Graycar, T. P., and Estell, D. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 5167−5171.
12. Wells, J. A., Powers, D. B., Bott, R. R., Graycar, T. P., and Estell, D. A. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1219−1223.
13. Wells, J. A., and Estell, D. A. (1988) *Trends Biochem. Sci. 13*, 291−297.
14. Bott, R., Ultsch, M., Wells, J. A., Powers, D., Burdick, D., Struble, M., Burnier, J., Estell, D., Miller, J., Graycar, T., Adams, R., and Power, S. (1987) *Biotech. Agric. Chem.* (Lebanon, H. M., Mumma, R. O., Honeycutt, R. C., and Duesing, J. H., Ed.) pp 139−147, ACS Symp. Ser. 334.
15. Russell, A. J., Thomas, P. G., and Fersht, A. R. (1987) *J. Mol. Biol. 193*, 803−813.
16. Ballinger, M. D., Tom, J., and Wells, J. A. (1996) *Biochemistry 33*, 13579−13585.
17. Rheinnecker, M., Eder, J., Pandey, P. S., and Fersht, A. R. (1994) *Biochemistry 33*, 221−225.
18. Rheinnecker, M., Baker, G., Eder, J., and Fersht, A. R. (1993) *Biochemistry 32*, 1199−1203.
19. Sørensen, S. B., Bech, L. M., Meldal, M., and Breddam, K. (1993) *Biochemistry 32*, 8994−8999.
20. Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., and Wells, J. A. (1986) *Science 233*, 659−663.

21. Takagi, H., Maeda, T., Ohtsu, I., Tsai, Y.-C., and Nakamori, S. (1996) *FEBS Lett. 395*, 127−132.
22. Takagi, H., Ohtsu, I., and Nakamori, S. (1997) *Protein Eng. 10*, 985−989.
23. Wangikar, P. P., Rich, J. O., Clark, D. S., and Dordick, J. S. (1995) *Biochemistry 34*, 12302−12310.
24. Bech, L. M., Sørensen, S. B., and Breddam, K. (1993) *Biochemistry 32*, 2845−2852.
25. Cornish, V. W., Mendel, D., and Schultz, P. G. (1995) *Angew. Chem., Int. Ed. Eng. 34*, 621−633.
26. Parsons, J. F., Xiao, G., Gilliland, G., and Armstrong, R. N. (1998) *Biochemistry 37*, 6286−6294.
27. Hohsaka, T., Ashizuka, Y., Murakami, H., and Sisido, M. (1996) *J. Am. Chem. Soc. 118*, 9778−9779.
28. Kuang, H., Brown, M. L., Davies, R. R., Young, E. C., and Distefano, M. D. (1996) *J. Am. Chem. Soc. 118*, 10702−10706.
29. Ory, J. J., Mazhary, A., Kuang, H., Davies, R. R., Distefano, M. D., and Branaszak, L. (1998) *Protein. Eng. 11*, 253−261.
30. Peterson, E. B., and Hilvert, D. (1995) *Biochemistry 34*, 6616−6620.
31. Suckling, C. J., and Zhu, L.-M. (1993) *Bioorg. Med. Chem. Lett. 3*, 531−534.
32. Rokita, S. E., and Kaiser, E. T. (1986) *J. Am. Chem. Soc. 108*, 4984−4987.
33. Kokubo, T., Sassa, S., and Kaiser, E. T. (1987) *J. Am. Chem. Soc. 109*, 606−607.
34. Radziejewski, C., Ballou, D. P., and Kaiser, E. T. (1985) *J. Am. Chem. Soc. 107*, 3352−3354.
35. Berglund, P., Stabile, M. R., Gold, M., Jones, J. B., Mitchinson, C., Bott, R. R., and Graycar, T. P. (1996) *Bioorg. Med. Chem. Lett. 6*, 2507−2512.
36. Berglund, P., DeSantis, G., Stabile, M. R., Shang, X., Gold, M., Bott, R. R., Graycar, T. P., Lau, T. H., Mitchinson, C., and Jones, J. B. (1997) *J. Am. Chem. Soc. 119*, 5265−5266.
37. DeSantis, G., Berglund, P., Stabile, M. R., Gold, M., and Jones, J. B. (1998) *Biochemistry 37*, 5968−5973.
38. DeSantis, G., and Jones, J. B. (1998) *J. Am. Chem. Soc. 120*, 8582−8586.
39. Plettner, E., Khumtaveeporn, K., Shang, X., and Jones, J. B. (1998) *Biorg. Med. Chem. Lett. 8*, 2291−2296.
40. Plettner, E., DeSantis, G., Stabile, M. R., and Jones, J. B. (1999) *J. Am. Chem. Soc. 121*, 4977−4984.
41. Kenyon, G. L., and Bruice, T. W. (1977) *Methods Enzymol. 47*, 407−430.
42. Wynn, R., and Richards, F. M. (1995) *Methods Enzymol. 251*, 351−356.
43. Brocklehurst, K. (1979) *Int. J. Biochem. 10*, 259−274.
44. Bech, L. M., and Breddam, K. (1988) *Carlsberg Res. Commun. 53*, 381−393.
45. Gloss, L. M., and Kirsch, J. F. (1995) *Biochemistry 34*, 12323−12332.
46. Smith, H. B., and Hartman, F. C. (1988) *J. Biol. Chem. 263*, 4921−4925.
47. Wynn, R., Harkins, P. C., Richards, F. M., and Fox, R. O. (1996) *Protein Sci. 5*, 1026−1031.
48. Foong, L. Y., You, S., Jaikaran, D. C. J., Zhang, Z., Zunic, V., and Woolley, G. A. (1996) *Biochemistry 36*, 1343−1348.
49. Holmgren, M., Liu, Y., Xu, Y., and Yellen, G. (1996) *Neuropharmacology 35*, 797−804.
50. Yang, N., George, A. L., and Horn, R. (1996) *Neuron 16*, 113−122.
51. Hubbell, W. L., Mchaourab, H. S., Altenbach, C., and Lietzow, M. A. (1996) *Structure 4*, 779−783.
52. Lin, Y., Nielsen, R., Murray, D., Hubbell, W. L., Mailer, C., Robinson, B. H., and Gelb, M. H. (1998) *Science 279*, 1925−1929.
53. Heinonen, P., Koskua, K., Pihlavisto, M., Marjamäki, A., Crockcroft, V., Savola, J.-M., Scheinin, M., and Lönnberg, H. (1998) *Bioconjugate Chem. 9*, 358−364.
54. Akabas, M. H., Kaufmann, C., Archdeacon, P., and Karlin, A. (1994) *Neuron 13*, 913−927.
55. Chen, J.-G., Liu-Chen, S., and Rudnick, G. (1997) *Biochemistry 36*, 1479−1486.
56. Shao, Z., and Arnold, F. H. (1996) *Curr. Opin. Struct. Biol. 6*, 513−518.
57. Lloyd, R., Dickman, M., and Jones, J. B. (1998) *Tetrahedron: Asymmetry 9*, 551−561.
58. van der Osten, C., Branner, S., Hastrup, S., Hedegaard, L., Rasmussen, M. D., Bisgard-Frantzen, H., Carlsen, S., and Mikkelson, J. M. (1993) *J. Biotechnol. 28*, 55−68.
59. Knapp, M., Daubermann, J., and Bott, R. R. Brookhaven Protein Data Bank, accession no. 1JEA.
60. Kuhn, P., Knapp, M., Soltis, S. M., Ganshaw, G., Thoene, M., and Bott, R. (1998) *Biochemistry 37* , 13446−13452.
61. Stabile, M. R., Lai, W. G., DeSantis, G., Gold, M., Jones, J. B., Mitchinson, C., Bott, R. R., Graycar, T. P., and Liu, C.-C. (1996) *Bioorg. Med. Chem. Lett. 6*, 2501−2506.
62. Grøn, H., Meldal, M., and Breddam, K. (1992) *Biochemistry 31*, 6011−6018.
63. Egmond, M. R., Antheunisse, W. P., van Bemmel, C. J., Ravestein, P., de Vlieg, J., Peters, H., and Branner, S. (1994) *Protein Eng. 7*, 793−800.
64. Maurer, K.-H., Markgraf, M., and Goddett, D. (1996) *Adv. Exp. Med. Biol. 379*, 243−256.
65. Olsen, O. H., Pedersen, J. T., Betzel, C., Eschenburg, S., Branner, S., and Hastrup S. (1996) *Adv. Exp. Med. Biol. 379*, 235−241.
66. Grøn, H., Bech, L. M., Branner, S., and Breddam, K. (1990) *Eur. J. Biochem. 194*, 897−901.
67. Fersht, A. (1985) *Enzyme Structure and Mechanism*, W. H. Freeman and Company, New York.
68. Bode, W., Meyer, E., and Powers, J. C. (1989) *Biochemistry 28*, 1951−1963.
69. Perona, J. J., Hedstrom, L., Rutter, W. J., and Fletterick, R. J. (1995) *Biochemistry 34*, 1489−1499.
70. Davis, B. G., Shang, X., DeSantis, G., and Bott, R. R. *Bioorg. Med. Chem.* (in press).
71. Hsia, C. Y., Ganshaw, G., Paech, C., and Murray, C. J. (1996) *Anal. Biochem. 242*, 221−227.
72. Bone, R., Frank, D., Kettner, C. A., and Agard, D. A. (1989) *Biochemistry 28*, 7600−7609.
73. Bone, R., Fujishige, A., Kettner, C. A., and Agard, D. A. (1991) *Biochemistry 30*, 10388−10398.
74. Bauer, C.-A., Brayer, G. D., Sielecki, A. R., and James, M. N. G. (1981) *Eur. J. Biochem. 120*, 289−294.
75. Mei, H.-C., Liaw, Y.-C., Li, Y.-C., Wang, D.-C., Takagi, H., and Tsai, Y.-C. (1998) *Protein Eng. 11*, 109−117.
76. Siezen, R. J., and Leunissen, J. A. M. (1997) *Protein Sci. 6*, 501−523.
77. Creemers, J. W. M., Siezen, R. J., Roebroek, A. J. M., Ayoubi, T. A. Y., Huylebroeck, D., and van de Ven, W. J. M. (1993) *J. Biol. Chem. 268*, 21826−21834.
78. Nakayama, K. (1997) *Biochem J. 327*, 625−635.
79. Bonneau, P. R., Graycar, T. P., Estell, D. A., and Jones, J. B. (1991) *J. Am. Chem. Soc. 113*, 1026−1030.
80. Wells, J. A. (1990) *Biochemistry 29*, 8509−8517.
81. $\Delta\Delta G^{\ddagger}T = RT \ln[(k_{cat}/K_M)_{WT}/(k_{cat}/K_M)_{CMM}]$; S166C-S-**b**, −2.04 kJ mol^{-1}; S166C-S-**d**, −4.21 kJ mol^{-1}; S166C-S-**h**, −1.54 kJ mol^{-1}; S166C-S-**i**, −3.20 kJ mol^{-1}; and S166C-S-**j**, −5.47 kJ mol^{-1}.
82. Nienaber, V. L., Breddam, K., and Birktoft, J. J. (1993) *Biochemistry 32*, 11469−11475.
83. Svendsen, I., Jensen, R. M., and Breddam, K. (1991) *FEBS Lett. 292*, 165−167.
84. Smyth, M. J., O'Connor, M. D., and Trapani, J. A. (1996) *J. Leukoc. Biol. 60*, 555−562.
85. Murphy, M. E. P., Moult, J., Bleackley, R. C., Gershenfeld, H., Weissman, I. L., and James, M. N. G. (1988) *Proteins: Struct., Funct., Genet. 4*, 190−204.
86. Caputo, A., James, M. N. G., Powers, J. C., Hudig, D., and Bleackley, R. C. (1994) *Nat. Struct. Biol. 1*, 364−367.
87. Ellman, G. L., Courtney, K. D., Andres Jr., V., and Featherstone, R. M. (1961) *Biochem. Pharmacol. 7*, 88−95.
88. *InsightII*, Biosym Technologies, Inc., San Diego, CA.
89. *Discover*, Biosym Technologies, Inc., San Diego, CA.

BI990861O