Pergamon

Tetrahedron: *Asymmetry* 10 (1999) 2563–2572

TETRAHEDRON: *ASYMMETRY*

Expanded structural and stereospecificity in peptide synthesis with chemically modified mutants of subtilisin

Kanjai Khumtaveeporn, Grace DeSantis and J. Bryan Jones *

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

Received 7 May 1999; accepted 10 June 1999

Abstract

Employing the strategy of combined site directed mutagenesis and chemical modification, we previously generated chemically modified mutant enzymes (CMMs) of subtilisin *Bacillus lentus* (SBL). We now report the use of these SBL-CMMs for peptide coupling reactions. The SBL-CMMs exhibit dramatically altered substrate specificity, including the acceptance of D-amino acid acyl donors, generating dipeptides containing D-Phe, D-Ala and D-Glu in up to 66% yield, which was not possible using wild-type SBL (WT-SBL). In addition, SBL-CMMs accommodate α-branched amino acids such as L-Ala-NH_2 as acyl acceptors in their S_1' pockets, which WT-SBL will not.

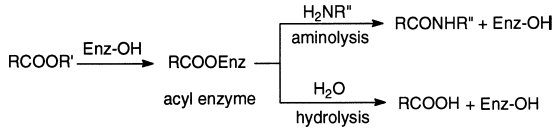
1. Introduction

Enzymatic, protease-catalyzed peptide coupling offers an attractive alternative to chemical solution or solid phase peptide synthesis.[1] Since enzymes are operational under mild reaction conditions and offer high regio- and stereoselectivity, the need for extensive protection strategies is minimized.[2] The serine proteases, which are a well-characterized group of hydrolytic enzymes, have demonstrated particular potential in peptide ligation.[1]

Serine protease-catalyzed peptide coupling requires the formation of a covalent acyl enzyme intermediate between the acyl donor substrate (RCOOR′) and the hydroxyl group of the catalytic serine residue, forming an acyl enzyme intermediate (RCOO-Enz) as outlined in Scheme 1. The acyl enzyme intermediate is then aminolyzed by the amino terminus of the acyl acceptor ($R''NH_2$), thereby effecting the formation of a peptide bond. However, in order to assure the success of this strategy the competing hydrolysis of the acyl enzyme intermediate, and also hydrolysis of the resulting peptide product (i.e. the enzyme's amidase activity), must be minimized. Furthermore, it is desirable to expand the substrate specificity of proteases in order to enhance their applicability to unnatural substrate analogues.

* Corresponding author. E-mail: jbjones@alchemy.chem.utoronto.ca

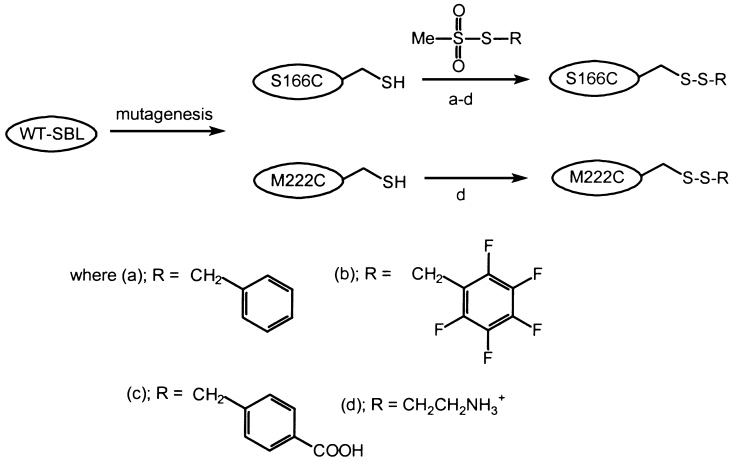
0957-4166/99/$ - see front matter
PII: S0957-4166(99)00255-4



Scheme 1.

Chemical modification,[3] genetic engineering,[4] the addition of water-miscible organic solvents,[5] enzyme immobilization, the use of high pH reaction conditions[6] and site directed mutagenesis[7] have already been explored with the goal of creating better peptide ligation biocatalysts. However, further progress is necessary since serine proteases have rather restricted P_1 and P_1' side-chain specificities, especially in their preference for the L-amino acid stereoisomers.[1] The incorporation of D-amino acids into peptides is of interest due to their prevalence in numerous biologically active compounds.[8] The current work is directed towards removing the above specificity limitations, with the aim of expanding the applicability of serine proteases for peptide coupling reactions.

Recently, we have successfully exploited the combination of site directed mutagenesis and chemical modification to modify the representative serine protease subtilisin *Bacillus lentus* (SBL).[9] This strategy involves the introduction of a cysteine residue at non-catalytic active site locations by site-directed mutagenesis. Since wild-type SBL (WT-SBL) does not contain any cysteine residues, the introduced cysteine provides a unique thiol for modification. The thiol group of the introduced cysteine is modified by reaction with a variety of methanethiosulfonate reagents to produce chemically modified mutant enzymes (CMMs). For this study, the S166C mutant, which is located as the bottom of the S_1 pocket and modulates P_1 specificity, and the M222C mutant, which is located at the entrance to the S_1' pocket and is adjacent to the catalytic triad residues, were subjected to chemical modification as outlined in Scheme 2.[10,11]



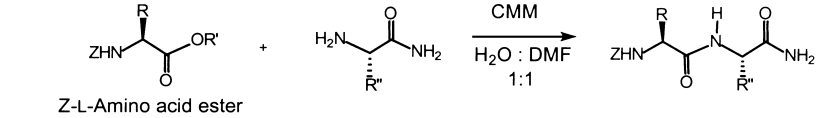
Scheme 2.

Several subtilisin CMMs (**a**–**d**) with significantly lowered amidase activity and which retain or surpass WT esterase activity levels were identified.[12,13] We now report the application of these selected CMMs as catalysts in peptide coupling reactions which demonstrate both increased yields and broadened specificity, in particular with respect to the incorporation of the unnatural D-amino acid series.

2. Results and discussion

The CMMs previously identified as having high esterase and lowered amidase activity[12,13] were selected as suitable candidate catalysts for peptide ligation. All of the CMMs (Scheme 2) chosen for the current peptide ligation study have esterase-to-amidase selectivity ratios up to 48-fold higher than WT.[12,13] Each of the CMMs were prepared and fully characterized as described previously.[12,13]

The peptide coupling reactions between acyl donors **1**–**4** and acyl acceptors **5** and **6** (Scheme 3) were conducted in 50% aqueous DMF. The acyl donors **1**–**4** were chosen as representative examples of large and small hydrophobic, negatively charged and positively charged P_1 side-chain-containing acyl donors, respectively, and permit the broad evaluation of the S_1 pocket specificity. The small amino acid amides **5** and **6** were chosen as the acyl acceptors since the S_1' pocket of subtilisin has narrow specificity and accepts α-branched amino acids only poorly.[14] The coupling reactions are shown in Scheme 3 and the results are summarized in Table 1.



Z-L-Amino acid ester

1, Z-L-Phe : R = CH_2Ph, R' = OBn
2, Z-L-Ala : R = CH_3, R' = OBn
3, Z-L-Glu : R = $(CH_2)_2COOH$, R' = OMe
4, Z-L-Lys : R = $(CH_2)_4NH_2$, R' = SBn

5, Gly-NH_2 : R" = H
6, L-Ala-NH_2 : R" = CH_3

7, Z-L-Phe-Gly-NH_2
8, Z-L-Ala-Gly-NH_2
9, Z-L-Glu-Gly-NH_2
10, Z-L-Lys-Gly-NH_2

11, Z-L-Phe-L-Ala-NH
12, Z-L-Ala-L-Ala-NH_2
13, Z-L-Glu-L-Ala-NH_2
14, Z-L-Lys-L-Ala-NH_2

Scheme 3.

The higher yield of Z-L-Phe-Gly-NH_2 (**7**) compared to Z-L-Ala-Gly-NH_2 (**8**), Z-L-Glu-Gly-NH_2 (**9**) and Z-L-Lys-Gly-NH_2 (**10**) with WT-SBL as the ligation catalyst (Table 1) attests to the preference of the S_1 pocket of SBL for large hydrophobic residues.[15] Furthermore, the yield of dipeptide **7** is excellent with each of the four S166-CMMs as a ligation catalyst, which is reflective of the high resistance of this pocket to attempts to change its specificity. Interestingly, the M222C-$SCH_2CH_2NH_3^+$ CMM, which has

Table 1
WT-SBL and SBL-CMM-catalyzed coupling of L-amino acids (**1**–**4**) and glycinamide (**5**)[a]

CMM	% Yield					
	Z-L-Phe-Gly-NH_2 (**7**)	Z-L-Ala-Gly-NH_2 (**8**)		Z-L-Glu-Gly-NH_2 (**9**)		Z-L-Lys-Gly-NH_2 (**10**)
	1 h	1 h	5 h	1 h	5 h	5 h
WT	92	68	91	64	62	83
S166C-SBn	92	59	95	68	96	93
S166C-$SCH_2C_6F_5$	93	42	94	61	61	71
S166C-SBnCOOH	quant.	38	82	30	62	99
S166C-$SCH_2CH_2NH_3^+$	95	63	94	69	quant.	86
M222C-$SCH_2CH_2NH_3^+$	33 (5 h)	-	-	-	33	-

[a] Reaction conditions : 0.1 mmol of acyl donor, 0.3 mmol of glycinamide hydrochloride, 0.4 mmol of Et_3N, 1 mg of enzyme, 1:1 H_2O : DMF. The total volume of the reaction is 1.0-1.2 mL.

Table 2
WT-SBL and SBL-CMM-catalyzed coupling of L-amino acids (**1**–**4**) and L-alaninamide (**6**)[a]

CMM	% Yield (24 h)			
	Z-L-Phe-Ala-NH_2 (**11**)	Z-L-Ala-Ala-NH_2 (**12**)	Z-L-Glu-Ala-NH_2 (**13**)	Z-L-Lys-Ala-NH_2 (**14**)
WT	57	0	0	0
S166C-SBn	51	0	0	0
S166C-$SCH_2C_6F_5$	33	0	0	0
S166C-SBnCOOH	48	0	0	0
S166C-$SCH_2CH_2NH_3^+$	88	16	14	0
M222C-$SCH_2CH_2NH_3^+$	22	0	0	0

[a] Conditions : 0.1 mmol of acyl donor, 0.2 mmol of alaninamide hydrochloride, 0.3 mmol of Et_3N, 1 mg of enzyme, 1:1 H_2O : DMF, 24h. The total volume of the reaction is 1.0-1.2 mL.

the highest esterase-to-amidase selectivity ratio of the CMMs evaluated here, gave only a 33% yield of dipeptide **7** after 5 h, and was accompanied by the recovery of a 41% yield of the starting material **1**.

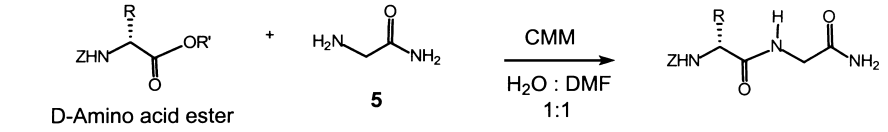
The yields reported in Table 1 illustrate that Z-L-Ala-OBn (**2**) and Z-L-Glu-OMe (**3**) are poorer acyl donors than Z-L-Phe-OBn (**1**). However, the yields of dipeptides Z-L-Ala-Gly-NH_2 (**8**) and Z-L-Glu-Gly-NH_2 (**9**) could generally be improved by increasing the reaction time from 1 h to 5 h. On the other hand, the yield of dipeptide Z-L-Glu-Gly-NH_2 (**9**) was not increased by longer reaction time when WT or S166C-$SCH_2C_6F_5$ was used as the ligation catalyst. Nevertheless, Z-L-Glu-Gly-NH_2 was isolated in quantitative yield after 5 h when S166C-S-$CH_2CH_2NH_3^+$, with its S_1 pocket possessing a positive charge complementary to the negative one of Glu-P_1 residue, was employed as the ligation catalyst, versus an only 62% yield with WT. The enhanced yield due to favorable S_1–P_1 electrostatic interaction is in line with our previous observation of rate enhancement of the S166C-$SCH_2CH_2NH_3^+$ CMM with the complementarily negatively charged P_1-Glu containing substrate Suc-AAPE-*p*-NA.[16] In all cases, M222C-$SCH_2CH_2NH_3^+$ proved to be a poor ligation catalyst.

Of the five CMMs evaluated, three CMMs gave an improved yield of the Z-L-Lys-Gly-NH_2 (**10**) dipeptide compared to WT. The almost quantitative yield of dipeptide **10** which results from the use of the complementary negatively charged S_1 pocket S166C-SBnCOOH ligation catalyst is attributed to a favorable electrostatic substrate–enzyme interaction, again consistent with previous kinetic experiments.[16] It is noteworthy that dipeptides **7**–**10** were isolated in good yield when S166C-SBn was used as the ligation catalyst. This is attributed to the high esterase-to-amidase ratio of this enzyme.

With the success of improved yields attributable to complementary P_1–S_1 electrostatic interaction in hand, the ability of the selected CMMs to accept α-branched amino acid acyl acceptors was investigated. The smallest α-branched amino acid, L-Ala-NH_2 (**6**) was used to probe the S_1' subsite tolerance. The constricted nature of the S_1' site was reflected by the longer reaction time of 24 h required when L-Ala-NH_2 was used as the acyl acceptor instead of Gly-NH_2 (**5**) (Table 2). In all cases, the reaction of the Z-L-Phe-OBn (**1**) acyl donor was slower with **6** than **5**. After 24 h only a 57% yield of dipeptide Z-L-Phe-L-Ala-NH_2 (**11**) was isolated with the WT ligation catalyst. The S166C-$SCH_2CH_2NH_3^+$ ligation catalyst gave improved dipeptide yields. In fact the excellent performance of this catalyst is exemplified by the fact that it is the only one which yields any Z-L-Ala-L-Ala-NH_2 (**12**) and Z-L-Glu-L-Ala-NH_2 (**13**) dipeptides. These results contrast the previously reported preference of the S_1' pocket of subtilisin *Bacillus lentus* for Ala over Gly.[15] In all cases yields obtained using Gly-NH_2 (**5**) as the acyl acceptor are higher, and the reaction faster (Table 1), compared to using L-Ala-NH_2 (**6**) (Table 2) as the acyl acceptor.

As the serine proteases, and WT-SBL in particular, have a normal preference for L-amino acid acyl

donors, the ability of the CMM approach to extend the applicability of SBL catalyzed peptide ligation to include D-amino acid ester (**15**–**18**) acyl donors (Scheme 4) was explored (Table 3). Excitingly, compared with the exclusive L-stereospecificity of the WT enzyme, the stereospecificity of all of the S166C-CMMs was broadened, resulting in good yields of D-amino acid containing dipeptides (**20**–**23**). While each of these enzymes still shows a preference for L-amino acids, yields of up to 66% of the D-amino acid containing dipeptide Z-D-Phe-Gly-NH_2 (**20**) were achieved using S166C-SBn as the ligation catalyst. This expansion of stereospecificity demonstrates a truly dramatic CMM-induced change in SBL's acceptance of D-amino acids.



15, Z-D-Phe : R = CH_2Ph, R' = OBn
16, Z-D-Ala : R = CH_3, R' = OBn
17, Z-D-Glu : R = $(CH_2)_2COOH$, R' = OMe
18, Z-D-Lys : R = $(CH_2)_4NH_2$, R' = OBn
19, Ac-D-Phe : R = CH_2Ph, R' = OBn, Z = Ac

20, Z-D-Phe-Gly-NH_2
21, Z-D-Ala-Gly-NH_2
22, Z-D-Glu-Gly-NH_2
23, Z-D-Lys-Gly-NH_2
24, Ac-D-Phe-Gly-NH_2, Z = Ac

Scheme 4.

The use of any of the CMM peptide ligation catalysts effected rather similar yields of the Z-D-Phe-Gly-NH_2 (**20**) and Z-D-Ala-Gly-NH_2 (**21**) dipeptides, suggesting that the mode of expansion of the stereospecificity is similar for all the CMMs. Since both large neutral, negatively and positively charged side chains were introduced into S_1 by chemical modification, and no discernible specificity pattern was observed, the mode of action of the observed stereospecificity relaxation is attributed to an altered binding mode of the D-amino acid acyl donors, such that carbobenzoxy (Z)-group becomes able to bind in the S_1 pocket.

When Z-D-Glu-OMe (**17**) was used as the acyl donor, 3 to 10% yields of dipeptide Z-D-Glu-Gly-NH_2 (**22**) were obtained with all CMM catalyzed reactions. While these yields are low, they represent a significant improvement over the zero yield with WT. Furthermore, replacing the Z-group of the Z-D-Phe-Gly-NH_2 (**15**) acyl donor with the acetyl group, as in Ac-D-Phe-Gly-NH_2 (**19**), resulted in lower yields (Table 3). This effect may be due to the differences in desolvation energies of the acetyl and Z-groups or to poorer binding of the acetyl group compared to the Z-group in the S_1 pocket.

To determine if any new beneficial interaction between the Z-D-Phe-OBn acyl donor and the S166C-

Table 3
WT-SBL and SBL-CMM-catalyzed coupling of D-amino acids (**15**–**19**) and glycinamide (**5**)[a]

CMM	% Yield (48h)				
	Z-D-Phe-Gly-NH_2 (**20**)	Z-D-Ala-Gly-NH_2 (**21**)	Z-D-Glu-Gly-NH_2 (**22**)	Z-D-Lys-Gly-NH_2 (**23**)	Ac-D-Phe-Gly-NH_2 (**24**)
WT	0	0	0	0	0
S166C-SBn	66	50	3	0	15
S166C-$SCH_2C_6F_5$	39	49	3	0	27
S166C-SBnCOOH	35	48	6	0	34
S166C-$SCH_2CH_2NH_3^+$	43	45	10	0	38

[a] Conditions: 0.1 mmol of acyl donor, 0.3 mmol of glycinamide hydrochloride, 0.4 mmol of Et_3N, 1 mg of enzyme, 1:1 H_2O : DMF, 48h. After 24h, another 1 mg of enzyme was added. The total volume of the reaction is 1.5-3.0 mL.

SBn CMM, which is not present in the WT, could be responsible for the enhanced reaction, molecular modeling analysis was employed. The energy of the structure of the ES complex of each of the Z-L-Phe-OBn and Z-D-Phe-OBn substrates bound to each of WT and S166C-SBn was minimized. The L-series of substrates were positioned in the enzyme active sites of WT and S166C-SBn such that the Z-group was oriented toward the S_2 pocket, the benzyl side chain of phenylalanine was oriented into the S_1 pocket and the -OBn leaving group of the ester was oriented toward the S_1' pocket. In contrast, the D-series of substrates were positioned in the enzyme active sites of WT and S166C-SBn such that the benzyl side chain of phenylalanine was oriented toward the S_2 pocket, the Z-group was oriented into the S_1 pocket and the -OBn leaving group of the ester was oriented toward the S_1' pocket. In the energy minimized structures, for three of the four structures the orientation of the P_1 and P_2 and P_1' oriented residues were quite similar. For the fourth ES complex between Z-D-Phe-OBn and S166C-SBn, the -OBn P_1' moiety was better oriented into the S_1' pocket than for the others, for which P_1' points toward S_2'. Since the S166C-SBn CMM shows the greatest expansion of stereospecificity with the Z-D-Phe-OBn acyl donor it is interesting that this ES pair adopts a different preferred binding mode to the others. However, the beneficial influence of the altered binding is not clear from the modeling.

3. Conclusion

The synthetic utility of chemically modified mutant enzymes for peptide synthesis has been demonstrated. Using the CMM approach we were able to incorporate a variety of amino acids as acyl donors into dipeptide products. These included charged P_1 amino acids which typically are not accepted into the S_1 pocket of WT-SBL. The specificity of the S_1' pocket of subtilisin was broadened to accept α-branched amino acids as efficient acyl acceptors. Moreover, significant broadening of S_1 pocket stereospecificity was also observed. Most dramatic is that all of the chemically modified mutant enzymes could accept D-amino acids as acyl donors while the WT enzyme accepted none. These results illustrate that the combined site-directed mutagenesis and chemical modification approach represents a powerful tool for broadening the use of enzymes in preparative organic synthesis applications, and further studies are in progress.

4. Experimental

4.1. General methods

WT-subtilisin *Bacillus lentus* and mutant enzymes, S166C and M222C were purified[9] and prepared as previously reported.[9c,13] Protected amino acids were purchased from Sigma or Bachem and were used as received. All solvents were reagent grade and distilled prior to use. Thin layer chromatography analysis and purification were performed on pre-coated Merck silica gel (60 F-254) plates (250 μm) visualized with UV light or iodine. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Gemini 200 (200 MHz for ^{1}H and 50.3 MHz for ^{13}C) or Unity 400 (400 MHz for ^{1}H and 100 MHz for ^{13}C) spectrometer and chemical shifts are given in ppm (δ) using $CDCl_3$ or DMSO-d_6 as an internal standard. High resolution mass spectra (HRMS) were recorded using Micromass ZAB-SE (FAB^+). Optical rotations were measured with a Perkin–Elmer 243B polarimeter.

4.2. General procedure for peptide ligation

To a solution of amino acid acyl donor (0.1 mmol) in DMF (0.4 mL) and water (0.4 mL), glycinamide hydrochloride (0.3 mmol) or alaninamide hydrochloride (0.2 mmol) and Et_3N (0.083–0.125 mL, 0.3–0.4 mmol) were added, followed by addition of a solution of 1 mg of active enzyme (0.0037 mmol, 0.037 equiv.), as determined by titration with phenylmethanesulfonyl fluoride (PMSF),[17] in buffer solution (10 mmol MES, 1 mmol $CaCl_2$, pH 5.8). The resulting total volume of reaction was 1.0–1.2 mL, with an apparent pH of 9.7. The reaction was left stirring at room temperature for the period of time indicated in Tables 1–3. In the case of D-amino acids being used as acyl donors, after 24 h, 1 mg more of active enzymes as well as an equal amount of DMF were added. After the reaction was finished, the mixture was then concentrated in vacuo and subjected to purification using preparative TLC (5–10% MeOH in CH_2Cl_2). Yields are given in Tables 1–3. The absence of background ligation was established by omitting enzyme from reaction mixture. Properties of the products are as follows.

4.2.1. *Z*-L-*Phe-Gly-NH$_2$* (**7**)[14a,18]

^{1}H NMR ($CDCl_3$) δ 3.10 (m, 2H, C*H*$_2$Ph), 3.85 (2×d, *J*=2, 5 Hz, 2H, NHC*H*$_2$CO), 4.40 (m, 1H, NHC*H*CO), 5.05 (s, 2H, OC*H*$_2$Ph), 5.50, 5.70, 6.25, 6.90 (4×br s, 4H, N*H*), 7.20–7.40 (m, 10H, 2×Ph); ^{13}C NMR ($CDCl_3$) δ 38.2, 42.7, 56.6, 67.3, 127.2, 128.1, 128.3, 128.6, 128.8, 129.2, 135.8, 136.0, 156.3, 171.2, 171.6. HRMS (FAB$^+$) MH$^+$ calcd for $C_{19}H_{22}N_3O_4$: 356.1610; found: 356.1613. $[\alpha]_D^{30}$=−4.3 (*c* 0.81, MeOH).

4.2.2. *Z*-L-*Ala-Gly-NH$_2$* (**8**)[19]

^{1}H NMR ($CDCl_3$) δ 1.40 (d, *J*=7 Hz, 3H, C*H*$_3$), 3.85 (dd, *J*=1.7, 5 Hz, 2H, NC*H*$_2$CO), 4.20 (m, 1H, NHC*H*CO), 5.10 (dd, *J*=1.6, 2 Hz, 2H, OC*H*$_2$Ph), 5.80, 6.60, 7.20 (3×br s, 4H, N*H*), 7.30–7.40 (m, 5H, Ph); ^{13}C NMR ($CDCl_3$) δ 17.7, 42.2, 50.6, 66.2, 127.6, 128.0, 136.0, 155.9, 171.3, 172.8. HRMS (FAB$^+$) MH$^+$ calcd for $C_{13}H_{18}N_3O_4$: 280.1297; found: 280.1310. $[\alpha]_D^{26}$=−8.44 (*c* 0.97, MeOH); lit. $[\alpha]_D^{23}$=−8.5 (*c* 2.0, MeOH).

4.2.3. *Z*-L-*Glu-Gly-NH$_2$* (**9**)[20]

^{1}H NMR (DMSO-d_6) δ 1.75–1.95 (m, 2H, C*H*$_2$CH$_2$COOH), 2.30 (m, 2H, CH$_2$C*H*$_2$COOH), 3.65 (dd, *J*=0.5, 1.7 Hz, 2H, NHC*H*$_2$CO), 4.10 (m, 1H, NHC*H*CO), 5.00 (s, 2H, OC*H*$_2$Ph), 7.20–7.40 (m, 5H, Ph), 12.40 (br s, 1H, COO*H*); ^{13}C NMR (DMSO-d_6) δ 27.2, 30.3, 41.9, 54.2, 65.6, 128.4, 128.5, 128.7, 128.8, 136.9, 156.2, 170.8, 171.7, 174.0. HRMS (FAB$^+$) MH$^+$ calcd for $C_{15}H_{20}N_3O_6$: 338.1352; found: 338.1332. $[\alpha]_D^{28}$=−10.2 (*c* 1.16, MeOH); lit. $[\alpha]_D^{25}$=−10.2 (*c* 1.0, MeOH).

4.2.4. *Z*-L-*Lys-Gly-NH$_2$* (**10**)

^{1}H NMR (DMSO-d_6) δ 1.50 (m, 2H, C*H*$_2$(CH$_2$)$_3$NH$_2$), 1.60–1.85 (m, 4H, CH$_2$C*H*$_2$C*H*$_2$CH$_2$NH$_2$), 2.00–2.18 (m, 2H, (CH$_2$)$_3$C*H*$_2$NH$_2$) 3.30 (m, 2H, NHC*H*$_2$CO), 4.40 (m, 1H, NHC*H*CO), 5.10 (s, 2H, OC*H*$_2$Ph), 6.05, 6.20 (2×br s, 2H, N*H*), 7.20–7.40 (m, 5H, Ph); ^{13}C NMR (DMSO-d_6) δ 28.1, 29.0, 32.2, 42.3, 45.9, 53.8, 66.7, 128.1, 128.6, 136.7, 155.6, 172.2, 175.3. HRMS (FAB$^+$) MH$^+$ calcd for $C_{16}H_{24}N_4O_4$: 337.1876; found: 337.1842. $[\alpha]_D^{28}$=+6.97 (*c* 0.55, MeOH).

4.2.5. *Z*-L-*Phe*-L-*Ala-NH$_2$* (**11**)[18,21]

^{1}H NMR (DMSO-d_6) δ 1.20 (d, *J*=7 Hz, 3H, CH$_3$), 2.90 (m, 2H, C*H*$_2$Ph), 4.30 (m, 2H, 2×NHC*H*CO), 4.90 (s, 2H, OC*H*$_2$Ph), 5.75, 6.10, 6.45 (3×br s, 4H, N*H*), 6.90–7.40 (m, 10H, 2×Ph); ^{13}C NMR (DMSO-d_6) δ 18.5, 37.4, 48.1, 56.2, 66.2, 126.3, 127.4, 127.7, 128.1, 128.3, 129.2, 137.1, 138.2, 155.9, 171.2,

174.1. HRMS (FAB$^+$) MH$^+$ calcd for $C_{20}H_{24}N_3O_4$: 370.1766; found: 370.1753. $[\alpha]_D^{29}$=−8.86 (*c* 0.57, MeOH).

*4.2.6. Z-*L*-Ala-*L*-Ala-NH$_2$ (**12**)*[22]

^{1}H NMR (DMSO-d_6) δ 1.20 (2×d, *J*=7 Hz, 6H, 2×CH$_3$), 4.10, 4.20 (m, 2H, 2×NHC*H*CO), 5.00 (s, 2H, OC*H*$_2$Ph), 7.20–7.40 (m, 5H, Ph); ^{13}C NMR (DMSO-d_6) δ 18.1, 18.4, 47.9, 50.1, 65.4, 127.7, 128.4, 137.0, 155.8, 172.0, 174.1. HRMS (FAB$^+$) MH$^+$ calcd for $C_{14}H_{20}N_3O_4$: 294.1454; found: 294.1457. $[\alpha]_D^{25}$=−20.4 (*c* 0.77, MeOH).

*4.2.7. Z-*L*-Glu-*L*-Ala-NH$_2$ (**13**)*

^{1}H NMR (DMSO-d_6) δ 1.20 (d, *J*=7 Hz, 3H, CH$_3$), 1.82–2.00 (m, 2H, C*H*$_2$CH$_2$COOH), 2.30 (m, 2H, CH$_2$C*H*$_2$COOH), 4.00, 4.20 (m, 2H, 2×NHC*H*CO), 5.00 (s, 2H, OC*H*$_2$Ph), 6.20–7.40 (m, 5H, Ph); ^{13}C NMR (DMSO-d_6) δ 18.4, 26.2, 30.2, 47.9, 53.1, 65.4, 126.5, 127.7, 127.8, 128.1, 128.4, 137.0, 156.2, 173.6, 173.8, 174.1. HRMS (FAB$^+$) MH$^+$ calcd for $C_{16}H_{22}N_3O_6$: 352.1509; found: 352.1478. $[\alpha]_D^{25}$=−16.7 (*c* 0.76, MeOH).

*4.2.8. Z-*D*-Phe-Gly-NH$_2$ (**20**)*

^{1}H and ^{13}C NMR data are identical to (**7**). HRMS (FAB$^+$) MH$^+$ calcd for $C_{19}H_{22}N_3O_4$: 356.1610; found: 356.1608; $[\alpha]_D^{30}$=+4.12 (*c* 1.17, MeOH).

*4.2.9. Z-*D*-Ala-Gly-NH$_2$ (**21**)*[23]

^{1}H and ^{13}C NMR data are identical to (**8**). HRMS (FAB$^+$) MH$^+$ calcd for $C_{13}H_{18}N_3O_4$: 280.1297; found: 280.1298; $[\alpha]_D^{27}$=+10.5 (*c* 0.72, MeOH); lit. $[\alpha]_D$=+10.5.

*4.2.10. Z-*D*-Glu-Gly-NH$_2$ (**22**)*

^{1}H and ^{13}C NMR data are identical to (**9**). HRMS (FAB$^+$) MH$^+$ calcd for $C_{15}H_{20}N_3O_6$: 338.1352; found: 338.1348; $[\alpha]_D^{28}$=+10.77 (*c* 1, MeOH).

*4.2.11. Ac-*D*-Phe-Gly-NH$_2$ (**24**)*[24]

^{1}H NMR (DMSO-d_6) δ 1.90 (s, 3H, CH$_3$), 3.05 (m, 2H, NHC*H*$_2$CO), 3.65 (2×d, *J*=7, 15 Hz, 2H, CHC*H*$_2$Ph), 4.40 (q, *J*=7 Hz, 1H, NHC*H*CO), 7.15–7.25 (m, 5H, Ph); ^{13}C NMR (DMSO-d_6) δ 22.4, 29.1, 45.9, 54.7, 126.1, 127.9, 128.9, 137.3, 155.2, 171.2, 171.5. HRMS (FAB$^+$) MH$^+$ calcd for $C_{13}H_{17}N_3O_3$: 264.1348; found: 264.1321. $[\alpha]_D^{30}$=−4.38 (*c* 0.80, MeOH).

4.3. Molecular modeling

The X-ray structure of subtilisin *Bacillus lentus*[25] 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.[26] 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)[25] 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,[27] on a Silicon Graphics Iris Indigo computer, using the consistent valence force field function (CVFF). A non-bonded cutoff distance

of 18 Å with a switching distance of 2 Å was employed. The non-bonded pair list was updated every 20 cycles and a dielectric constant of 1 was used in all calculations. The energy of the structure of 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 were minimized, and then all of the atoms were minimized. The AAPF product inhibitor was subsequently replaced with the appropriate acyl donor substrate Z-L/D-Phe-OBn and then the energy of this structure was minimized as above.

Acknowledgements

We are grateful to the National Sciences and Engineering Research Council of Canada and Genencor International Inc. for generous funding, and Genencor for providing WT-, S166C- and M222C-SBL enzymes. We also thank Dr. Erika Plettner for preliminary experiments and Dr. Benjamin G. Davis for helpful discussion.

References

1. (a) Wong, C.-H.; Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*; Pergamon Press: New York, 1994; pp. 46–59. (b) Faber, K. *Biotransformations in Organic Chemistry*, 2nd ed.; Springer-Verlag: Heidelberg, 1995; pp. 298–305. (c) Jakubke, H.-D.; Kuhl, P.; Konnecke, A. *Angew. Chem., Int. Ed. Engl*. **1985**, *24*, 85–93. (d) For a comprehensive Protease-Catalyzed Peptide Synthesis Database up to 1993 see: http://biont.ukc.ac.uk/bio/database/proteases/Entry.html
2. Sears, P.; Wong, C.-H. *Biotechnol. Progress* **1996**, *12*, 423–433, and references cited therein.
3. (a) Bech, L. M.; Breddam, K. *Carlsberg Res. Commun*. **1988**, *53*, 381–393. (b) Smith, H. B.; Hartman, F. C. *J. Biol. Chem*. **1988**, *263*, 4921–4925. (c) Gron, H.; Bech, L. M.; Branner, S.; Breddam, K. *Eur. J. Biochem*. **1990**, *194*, 897–901. (d) Wynn, R.; Richards, F. M. *Protein Sci*. **1993**, *2*, 395–403. (e) Gloss, L. M.; Kirsch, J. F. *Biochemistry* **1995**, *34*, 12323–12332. (f) Wynn, R.; Harkins, P. C.; Richards, F. M.; Fox, R. O. *Protein Sci*. **1996**, *5*, 1026–1031.
4. (a) Khosla, C.; Caren, R.; Kao, C. M.; McDaniel, R.; Wang, S.-W. *Biotechnol. Bioeng*. **1996**, *52*, 122–128. (b) Shao, Z.; Arnold, F. H. *Curr. Opin. Struct. Biol*. **1996**, *6*, 513–518.
5. (a) Barbas III, C. F.; Matos, J. R.; West, B. J.; Wong, C.-H. *J. Am. Chem. Soc*. **1988**, *110*, 5162–5166. (b) Stepanov, V. M. *Pure Appl. Chem*. **1996**, *68*, 1335–1339. (c) Wong, C.-H.; Chen, S.-T.; Hennen, W. J.; Bibb, J. A.; Wang, Y.-F.; Liu, J. L.-C.; Pantoliano, M. W.; Whitlow, M.; Bryan, P. N. *J. Am. Chem. Soc*. **1990**, *112*, 945–953.
6. Barbas III, C. F.; Wong, C.-H. *J. Chem. Soc., Chem. Commun*. **1987**, 533–534.
7. (a) Bonneau, P. R.; Graycar, T. P.; Estell, D. A.; Jones, J. B. *J. Am. Chem. Soc*. **1991**, *113*, 1026–1030. (b) Graycar, T. P.; Bott, R. R.; Caldwell, R. M.; Daubermann, J. L.; Lad, P. J.; Power, S. D.; Sagar, H. I.; Silva, R. A.; Weiss, G. L.; Woodhouse, L. R.; Estell, D. A. *Ann. N.Y. Acad. Sci*. **1992**, *672*, 71–79. (c) Abrahmsen, L.; Tom, J.; Burnier, J.; Butcher, K. A.; Kossiakoff, A.; Wells, J. A. *Biochemistry* **1991**, *30*, 4151–4159.
8. (a) Margolin, A. L.; Tai, D.-F.; Klibanov, A. M. *J. Am. Chem. Soc*. **1987**, *109*, 7885–7887. (b) Chen, S.-T.; Chen, S.-Y.; Wang, K.-T. *J. Org. Chem*. **1992**, *57*, 6960–6995. (c) West, J.; Wong, C.-H. *J. Org. Chem*. **1986**, *51*, 2728–2735. (d) West, J. B.; Scholten, J.; Stolowich, N. J.; Hogg, J. L.; Scott, A. I.; Wong, C.-H. *J. Am. Chem. Soc*. **1988**, *110*, 3709–3710. (e) Jonsson, A.; Wehtje, E.; Adlercreutz, P.; Mattiason, B. *J. Mol. Catal. B* **1996**, 43–51. (f) Kawashiro, K.; Sugahara, H.; Tsukioka, T.; Sugiyama, S.; Hayashi, H. *Biotechnol. Lett*. **1996**, *18*, 1381–1386. (g) Chen, S. T.; Chen, S. Y.; Tu, C. C.; Chiou, S. H.; Wang, K.-T. *J. Protein Chem*. **1995**, *14*, 205–215. (h) Chen, S. T.; Chen, S. Y.; Chen, H. J.; Huang, H. C.; Wang, K. T. *Bioorg. Med. Chem. Lett*. **1993**, *3*, 727–733. (i) So, J.-E.; Kang, S.-H.; Kim, B.-G. *Enzyme Microb. Technol*. **1998**, *23*, 211–215.
9. (a) Stabile, M. R.; Lai, G. W.; DeSantis, G.; Gold, M.; Jones, J. B.; Mitchinson, C.; Bott, R. R.; Graycar, T. P.; Liu, C.-C. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 2501–2506. (b) Berglund, P.; Stabile, M. R.; Gold, M.; Jones, J. B.; Mitchinson, C.; Bott R. R.; Graycar, T. P. *Bioorg. Med. Chem. Lett*. **1996**, *6*, 2507–2512. (c) DeSantis, G.; Berglund, P.; Stabile, M. R.; Gold, M.; Jones, J. B. *Biochemistry* **1998**, *37*, 5968–5973. (d) Berglund, P.; DeSantis, G.; Stabile, M. R.; Shang, X.; Gold, M.; Bott, R. R.; Graycar, T. P.; Lau, T. H.; Mitchinson, C.; Jones, J. B. *J. Am. Chem. Soc*. **1997**, *119*, 5265–5266.

10. Nomenclature of Schechter, I.; Berger, A. *Biochem. Biophy. Res. Commun*. **1967**, *27*, 157–162.
11. Bott, R. R.; Ultsch, M.; Kossiakoff, A.; Graycar, T. P.; Katz, B.; Power, S. *J. Biol. Chem*. **1988**, *263*, 7895–7906.
12. Plettner, E.; Khumtaveeporn, K.; Shang, X.; Jones, J. B. *Bioorg. Med. Chem. Lett*. **1998**, *8*, 2291–2296.
13. Plettner, E.; DeSantis, G.; Stabile, M. R.; Jones, J. B. *J. Am. Chem. Soc*. **1999**, *121*, 4977–4981.
14. (a) Moree, W. J.; Sears, P.; Kawashiro, K.; Witte, K.; Wong, C.-H. *J. Am. Chem. Soc*. **1997**, *119*, 3942–3947. (b) Betzel, C.; Klupsch, S.; Padendorf, G.; Hastrup, S.; Branner, S.; Wilson, K. S. *J. Mol. Biol*. **1992**, *223*, 427–445. (c) Sears, P.; Schuster, M.; Wang, P.; Witte, K.; Wong, C.-H. *J. Am. Chem. Soc*. **1994**, *116*, 6521–6530. (d) Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. *Science* **1994**, *266*, 243–247.
15. Gron, H.; Meldal, M.; Breddam, K. *Biochemistry* **1992**, *31*, 6011–6018.
16. DeSantis, G.; Shang, X.; Jones, J. B. *Biochemistry* **1999**, *38*, in press.
17. Hsia, C. Y.; Gunshaw, G.; Paech, C.; Murray, C. J. *Anal. Biochem*. **1996**, *242*, 221–227.
18. Morihara, K.; Oka, T. *Biochem. J*. **1977**, *163*, 531–542.
19. Bodanszky, M.; Bodanszky, A.; Casaretto, M.; Zahn, H. *Int. J. Peptide Protein Res*. **1985**, *26*, 550–556.
20. Schon, I.; Szirtes, T.; Uberhardt, T.; Rill, A.; Csehi, A.; Hegedus, B. *Int. J. Peptide Protein Res*. **1983**, *22*, 92–109.
21. Brubacher, J.; Zaher, M. R. *Can J. Biochem*. **1979**, *57*, 1054–1072.
22. Katakai, R.; Oya, M.; Toda, F.; Uno, K.; Iwakura, Y. *Macromolecules* **1973**, *6*, 827–831.
23. Richman, S. J.; Goodman, M.; Nguyen, T. M.-D.; Schiller, P. W. *Int. J. Peptide Protein Res*. **1985**, *25*, 648–662.
24. Thompson, S. A.; Andrew, P. R.; Hanzlik, R. P. *J. Med. Chem*. **1986**, *29*, 104–111.
25. Knapp, M.; Daubermann, J.; Bott, R. R. *Brookhaven Database Entry 1JEA*.
26. *Insight II*, Version 2.3.0 Biosym Technologies, Inc. San Diego, CA, USA.
27. *Discover*, Version 2.9.5 Biosym Technologies, Inc. San Diego, CA, USA.