Chemically Modified “Polar Patch” Mutants of Subtilisin in Peptide Synthesis with Remarkably Broad Substrate Acceptance: Designing Combinatorial Biocatalysts

Kazutsugu Matsumoto,[b] Benjamin G. Davis,*[a] and J. Bryan Jones*[c]

Abstract: A significant enhancement of the applicability of the serine protease subtilisin Bacillus lentus (SBL) in peptide synthesis was achieved by using the strategy of combined site-directed mutagenesis and chemical modification to create chemically modified mutant (CMM) enzymes. The introduction of polar and/or homochiral auxiliary substituents, such as X/C136 oxazolidinones, alkylammonium groups, and carbohydrates at position 166 at the base of the primary specificity S1 pocket created SBL CMMs S166C-S-X with strikingly broad structural substrate specificities. These CMMs are capable of catalyzing the coupling reactions of not only L-amino acid esters but also d-amino acid esters as acyl donors with glycaminamide to give the corresponding dipeptides in good yields. These powerful enzymes are also applicable to the coupling of L-amino acid acyl donors with α-branched acyl acceptor, l-alaninamide. Typical increases in isolated yields of dipeptides of 60–80% over SBL-WT (e.g., 0% yield of Z-d-Glu-GlyNH₂ using SBL-WT → 74% using S166C-S-(CH₂)₃ NMe₃⁺) demonstrate the remarkable synthetic utility of this “polar patch” strategy. Such wide-ranging systems displaying broadened and therefore similarly high, balanced yields of products (e.g., 91% Z-l-Ala-GlyNH₂ and 86% yield of Z-d-Ala-GlyNH₂ using S166C-S-(3R,4S)-indenooxazolidinone) may now allow the use of biocatalysts in parallel library synthesis.

Keywords: biotransformations · enantioselectivity · hydrolases · mutagenesis · protein modifications

Introduction

Enzymatic peptide coupling is an attractive method for the preparation of a variety of peptides because this method requires minimal protection of the substrate, proceeds under mild conditions and without racemization.[1] In spite of these advantages, two major problems can limit the use of serine proteases in peptide synthesis. One is their efficient proteolytic (amidase) activity which causes hydrolysis of the coupled peptide product, and the other is their stringent structural and stereospecificity which typically confines their use in synthesis to a limited range of L-amino acid substrates (Scheme 1).

Since the nature of the amino acids or peptides to be coupled may be determined in synthesis simply through appropriate choice of donor and acceptor, it is largely the efficiency, mildness and environmentally benign nature of a given biocatalyst that the chemist wishes to exploit. In essence, it would be highly advantageous to discard the substrate specificity while retaining the reactivity benefits listed above. An area of growing interest is that of combinatorial biocatalysis,[2, 3] the use of enzyme catalysts in parallel reactions to provide arrays of related molecules. However, although combinatorial chemistry has revolutionised the approach to traditional chemical synthesis, the development of combinatorial biocatalysis has been hampered by the often-
stringent substrate specificities of synthetically useful enzymes. As a solution to these problems we have set ourselves the goal of creating biocatalysts with the ability to process a broad range of substrates through the engineering of new broad substrate specificities, while maintaining catalytic efficiency and thereby expanding synthetic utility.

Recently, the combination of site-directed mutagenesis and chemical modification has been recognized as a powerful technique for the efficient and rapid creation of new active-site environments in enzymes. We have previously exploited this method to improve enzyme activity and alter substrate specificity of the subtilisin from Bacillus lenticus (SBL, EC 3.4.21.62 formerly EC 3.4.21.14). Herein, we present a significant enhancement of the applicability of chemically modified mutants of SBL (CMMs) as catalysts for peptide couplings that can incorporate D-amino acid esters as acyl donors and an α-branched amino acid amide as an acyl acceptor, neither of which is possible with SBL-wild type (WT). This has allowed the use of these novel biocatalysts in parallel reaction arrays.

**Results and Discussion**

Strategies for enhancing synthetic utility and broadening specificity: Several strategies have been employed in an attempt to overcome specificity restrictions. As alternative catalysts with typically broader specificities, lipases have been employed but their lower activities in aqueous solvents, negates a broader substrate specificity with a disadvantageous solvent specificity. Alternatively, elegant substrate engineering or mimcry has been employed but by definition requires the synthesis of specifically designed substrates that typically are not readily available. Impressive protein engineering methods, such as site directed mutagenesis or forced evolution have also successfully altered the specificities of proteases for use in synthesis. However, these methods do not routinely allow the creation of large numbers of catalysts in sufficient quantities for synthesis.

To overcome these problems, biocatalysts with broad specificities for peptide synthesis that are readily created, that may be used in simple solvent systems and that employ standard readily available substrates are required. Using the serine protease subtilisin Bacillus lenticus (SBL), we chose a combined site-directed mutagenesis and chemical modification approach. This strategy involves the introduction of a cysteine residue at a key active-site position through site-directed mutagenesis and the reaction of its thiol side chain with highly chemoselective methanethiosulfonate (MTS) reagents to produce chemically modified mutant (CMM) enzymes (Scheme 2). Since wild-type (WT) SBL contains no natural cysteines the site of modification corresponds only to the site of mutagenesis. The ease of this method is such that large arrays of different CMM biocatalysts can be created and screened in a short space of time.

Using this technique polar groups intended to influence the stereochemical P₁ specificity of such ligations. Carbohydrates and oxazolidinones are used widely as effective chiral auxiliaries and we wished to probe the effect of introducing the homochiral groups into the primary binding region of an existing chiral enzyme active site environment as a tactic for broadening the stereochemical tolerance of substrate specificity. We envisaged that this might increase the potential for stereochemical mismatching with a given chiral substrate thereby reducing the difference in energy between the diastereomeric transition states for preferred (e.g., L-amino acids) and non-preferred (e.g., D-amino acids) substrates. Furthermore, we envisaged that the use of such polar groups, all of which increase the number of potential hydrogen bond donors and acceptors at position 166 within the typically hydrophobic S₁ pocket, would also dramatically affect the hydrophobic vs hydrophilic specificity profile.

The construction of CMMs bearing carbohydrates (glyco-CMMs) as chiral auxiliaries in their binding pockets, S166C-S-a-d: We selected and constructed a range of glycosylated SBL serine proteases (glyco-CMMs) in which representative carbohydrates a-d (see Scheme 2) were attached to the interior of the primary S₁ binding pocket by selective glycosylation at position 166. Residue S166 was selected for mutagenesis and modification because it is located at the bottom of the S₁ pocket which modulates P₁ substrate specificity (Figure 1) and is therefore the pocket that determines acyl donor selectivity in amino acid ester peptide
Table 1. Kinetic parameters of WT and CMMs of SBL.

<table>
<thead>
<tr>
<th>Kinetic parameter</th>
<th>SBL-WT</th>
<th>S166C-S-a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_{cat}/K_{M})amide</td>
<td>3593</td>
<td>5058</td>
<td>4264</td>
<td>6033</td>
<td>3241</td>
<td>1495</td>
<td>3277</td>
<td>1488</td>
<td>4556</td>
<td>4898</td>
<td>2400</td>
<td>1800</td>
<td>4900</td>
</tr>
<tr>
<td>(k_{cat}/K_{M})ester</td>
<td>209</td>
<td>109</td>
<td>112</td>
<td>81</td>
<td>58</td>
<td>22</td>
<td>52</td>
<td>20</td>
<td>47</td>
<td>75</td>
<td>23</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>E/A</td>
<td>17</td>
<td>46</td>
<td>38</td>
<td>75</td>
<td>56</td>
<td>68</td>
<td>63</td>
<td>74</td>
<td>97</td>
<td>65</td>
<td>104</td>
<td>225</td>
<td>350</td>
</tr>
</tbody>
</table>

[a] Kinetic constants determined in duplicate using the low substrate concentration approximation in 0.1 M Tris buffer, pH 8.6, 0.005 % Tween 80, 1 % DMSO with succ-AAPF-SBn as substrate. [b] Michaelis–Menten constants were measured at 25 °C according to the initial rates method in 0.1 M Tris-HCl buffer at pH 8.6, 0.005 % Tween 80, 1 % DMSO, succ-AAPF-pNA as the substrate.

Use of carbohydrate-modified CMMs in peptide ligations: Firstly, we probed the structural breadth of the P specificity of S166C-S-a-d by examining the ligations of l-amino acids, Z-l-PheOBn (2), Z-l-AlaOBn (3), and Z-l-GluOMe (4) as acyl donors, with glycineamide (8) as an acyl acceptor (Table 2). These reactions were carried out by using 0.5 mg [28] of an active enzyme in a simple 1:1 water:DMF solution system. In all cases, the reactions proceeded smoothly to afford the corresponding dipeptides 10–12 in good yields and were complete within 5 h or less. In accord with our goals of not reducing the inherent substrate breadth of SBL, these results indicated that the introduction of groups a–d did not affect the essential ability of SBL to accept l-amino acids as acyl acceptors in peptide coupling. Good yields of Z-l-Phe-
Reactions were performed in DMF/water (1/1, v/v) at a concentration of 0.1 mM acyl donor, 0.3 mM acyl acceptor, and 0.3 mM Et$_3$N in the presence of 1 mg mL$^{-1}$ of active enzyme at room temperature unless otherwise noted. Under the same conditions, spontaneous hydrolysis or aminolysis did not occur. In these cases, slightly reduced yields of Z-$\text{amino acid}$-$\text{amino acid}$ (8 %) were observed (the highest yield of Z-$\text{amino acid}$-$\text{amino acid}$ was 8 % by using S166C-S-$\text{amino acid}$-$\text{amino acid}$). However, even this, glyco-CMMs S166C-S-$\text{amino acid}$-$\text{amino acid}$ were catalysts for the coupling of $\text{amino acid}$-$\text{amino acid}$ with $\text{amino acid}$ as acyl acceptor, and 0.3 mM Et$_3$N in the presence of 1 mg mL$^{-1}$ of active enzyme added and the mixture was stirred for another 24 h.

<table>
<thead>
<tr>
<th>Acyl donor</th>
<th>Acyl acceptor</th>
<th>Product</th>
<th>Time [h]</th>
<th>Isolated yield [%] with S166C-S-$\text{amino acid}$-$\text{amino acid}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z-$\text{L-PheOBn}$ (2)</td>
<td>Gly-NH$_2$ · HCl (8)</td>
<td>Z-$\text{L-PheGlyNH}_2$ (10)</td>
<td>1</td>
<td>92 ± 64</td>
</tr>
<tr>
<td>Z-$\text{L-AlaOBn}$ (3)</td>
<td>8</td>
<td>Z-$\text{L-AlaGlyNH}_2$ (11)</td>
<td>5</td>
<td>91 ± 85</td>
</tr>
<tr>
<td>Z-$\text{L-GluOMe}$ (4)</td>
<td>8</td>
<td>Z-$\text{L-GluGlyNH}_2$ (12)</td>
<td>62 ± 58</td>
<td>54 ± 67</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>t-$\text{AlaNH}_2$ · HCl (9)</td>
<td>24$^{[9]}$</td>
<td>57 ± 28</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>Z-$\text{L-Ala-L AlaNH}_2$ (13)</td>
<td>24$^{[9]}$</td>
<td>0 ± 15</td>
</tr>
<tr>
<td>Z-$\text{d-PheOBn}$ (5)</td>
<td>8</td>
<td>Z-$\text{d-PheGlyNH}_2$ (16)</td>
<td>24$^{[9]}$</td>
<td>0 ± 48</td>
</tr>
<tr>
<td>Z-$\text{d-AlaOBn}$ (6)</td>
<td>8</td>
<td>Z-$\text{d-AlaGlyNH}_2$ (17)</td>
<td>48$^{[9]}$</td>
<td>0 ± 80</td>
</tr>
<tr>
<td>Z-$\text{d-GluOBn}$ (7)</td>
<td>8</td>
<td>Z-$\text{d-GluGlyNH}_2$ (18)</td>
<td>48$^{[9]}$</td>
<td>0 ± 63</td>
</tr>
</tbody>
</table>

[a] Reactions were performed in DMF/water (1/1, v/v) at a concentration of 0.1 mM acyl donor, 0.3 mM acyl acceptor, and 0.3 mM Et$_3$N in the presence of 1 mg mL$^{-1}$ of active enzyme at room temperature unless otherwise noted. Under the same conditions, spontaneous hydrolysis or aminolysis did not occur. [b] In these cases, slightly reduced yields of Z-$\text{amino acid}$-$\text{amino acid}$ (8 %) were observed (the highest yield of Z-$\text{amino acid}$-$\text{amino acid}$ was 8 % by using S166C-S-$\text{amino acid}$-$\text{amino acid}$). However, even this, glyco-CMMs S166C-S-$\text{amino acid}$-$\text{amino acid}$ were catalysts for the coupling of $\text{amino acid}$-$\text{amino acid}$ with $\text{amino acid}$ as acyl acceptor, and 0.3 mM Et$_3$N in the presence of 1 mg mL$^{-1}$ of active enzyme added and the mixture was stirred for another 24 h.

The construction of CMMs bearing oxazolidinones as chiral auxiliaries in their binding pockets, S166C-S-e-b. In the light of these early exciting indications obtained through the introduction of homochiral substituents to the S$_i$ binding pocket of SBL, we wished to probe the nature of this useful strategy. We recently introduced our work in this area. The construction of CMMs bearing oxazolidinones as chiral auxiliaries in their binding pockets, S166C-S-e-b. In the light of these early exciting indications obtained through the introduction of homochiral substituents to the S$_i$ binding pocket of SBL, we wished to probe the nature of this useful strategy. We recently introduced our work in this area.
these differences and to thoroughly investigate the possibility of mismatched pairing within the transitions states of these peptide ligation reactions we chose to introduce enantiomeric substituents to position 166. t-galactose and t-glucose, which would be required for the synthesis of the enantiomers of 1a–d, are not readily available chironns and we chose instead to create two enantiomeric pairs of substituents based on the well-established oxazolidinone auxiliaries. S166C was modified with 1e–h to create the diastereomeric oxazolidinonyl-CMMs S166C-e–h in essentially an analogous manner to that used for the construction of glyco-CMMs S166C-a–d.

Again we were delighted to observe that all oxazolidinonyl-CMMs S166C-e–h displayed enhanced esterase to amidase activity ratios (E/A = 63–97) over SBL-WT (Table 1). Excitingly, striking differences in rates were observed for CMMs modified with oxazolidinones of opposite absolute configuration (S166C-S cf. f and S166C-S-g cf. -h) in the hydrolysis of chiral substrates suc-AAPF-pNA and suc-AAPF-SBn and this indicated the possibility of pair matching with t-Phe-containing substrates. For example, (k, k')ex and (l, l')ex of S166C-S-g bearing the R,S-indenooxazolidinonyl side chain -g at position 166 is three-fold greater than that of S166C-S-h bearing the S,R-indenooxazolidinonyl side chain -h.

Use of oxazolidinone-modified CMMs in peptide ligations: As for S166C-a–d, we initially probed the P1 specificity of S166C-S-e–h by examining the ligation of t-amino acids, Z-t-PheOBn (2), Z-t-AlaOBn (3), and Z-t-GluOMe (4) as acyl acceptors, with glycineamide (8) as an acyl acceptor (Table 2). Again, as for S166C-a–d, useful to excellent yields of 10 (82–100%), 11 (87–95%) and 12 (54–68%) were obtained. Similarly, better catalysts than SBL-WT were obtained for each ligation; for example, S166C-S-h for 10 and 11, S166C-S-h for 12. Notably, the markedly different (k, k')ex values of S166C-S-g and -h with substrates containing t-Phe as the P1 substituent (see above, Table 1), were partially reflected in the yields of Z-t-PheGly-NH2 (10) obtained from their use as ligation catalysts: 82% using S166C-S-g yet 100% from S166C-S-h under identical conditions. Moreover, reverse yield differences were observed with these same two catalysts in the synthesis of Z-t-GluGly-NH2 (12) with a S166C-S-g-catalyzed reaction yield of 68% being superior in that reaction by 12% to that in the S166C-S-h-catalyzed reaction.

In contrast, little yield difference was observed in the secondary effects of the enantiomeric side chains e versus f and g versus h in the S1 pocket upon P1 specificity in liganations of t-alaninamide (8). However, despite this lack of variation the enhanced utility of oxazolidinonyl-CMMs S166C-S-e–h over SBL-WT was still demonstrated by the ability of them all to synthesize C-terminal t-alaninyl dipeptides 13 and 14 and 15 in yields of up to 37, 21 and 61%, respectively.

Oxazolidinonyl-CMMs S166C-S-e–h proved to be amongst the most effective t-amino acid-acylating CMM catalysts. For example, the use of S166C-S-h in the synthesis of 16 increased the yield by two-fold over glyco-CMMs S166C-a–d to 14%; again this is an immeasurable enhancement over the inability of SBL-WT to catalyse this reaction. Similarly, S166C-e and -g both catalyzed the synthesis of Z-t-AlaGly-NH2 (17) in very good yields of 88%, and are very close to the best reported yields for any enzyme catalyzed C-terminal t-amino acid ligation. Again, it should be noted that all of S166C-S-e–h were able to catalyse the syntheses of 16–18, whereas SBL-WT cannot. As for the ligations of t-amino acids with 8, these ligations of t-amino acids with 8 showed some variations according to the configuration of the oxazolidinonyl substituent. For example, S166C-S-h bearing the S,b-indenooxazolidinonyl side chain at position 166 gave a two-fold higher yield (14%) than S166C-S-g bearing the R,S-indenooxazolidinonyl side chain -g (7% yield). Although in some cases these variations were only slight, the unambiguous nature of the installation of such enantiomeric side chains allows comparison with a high degree of precision.

The construction of CMMs bearing achiral groups in their binding pockets, S166C-S-i–l: Although such noticeble variations according to configuration were observed in the yields obtained with CMMs bearing different homochiral side chains at position 166, it was clear that such diastereomeric effects could not alone account for the very substantial broadening in the stereospecificity of S166C-S-a–h as illustrated by increases in yield as large as 88% over SBL-WT (e.g., an increase of 0% → 88% yield of Z-t-AlaGlyNH2 (17) by changing the catalyst from SBL-WT → S166C-S-g). In order to dissect the origin of the observed enhancements in synthetic utility, four further achiral side chains were introduced at position 166 to create CMMs S166C-S-i–l (Scheme 2).

Firstly, S166C-S-i, containing an unsubstituted oxazolidinonyl framework at position 166 was constructed. In addition, S166C-S-j,k bearing simple, highly polar side chain functionalities: singly charged -SCH2CH2NMe3+ and triply charged -SCH2(C2H4NH3)+3, respectively were constructed to examine the effect of simply introducing different levels of charge into the S1 pocket. Finally, non-polar neohexyl side chain -SCH2CH2CMe3 (4), which is near isosteric to side chain -j, was introduced to create S166C-S-i and so provide a rough estimate of the effect of polarity in S166C-S-j when corrected for underlying steric and hydrophobic effects. Again enhanced E/A values over SBL-WT were observed for all of S166C-S-i–l (Table 1).

Use of CMMs modified with polar achiral groups in peptide ligations: Although, the use of CMM S166C-S-i, bearing an achiral oxazolidinonyl side chain at position 166, as a catalyst resulted in slight reductions in yield as compared with those with a chiral oxazolidinonyl side chain CMMs S166C-S-e–h for some syntheses (11, 14, 17; for example by 19–27% for 17), the majority of yields were either similar (10, 12, 16, 18) or indeed improved (13, 15). Thus, it seems that although diastereomeric effects may be observed in certain selected examples, the major role of side chains a–i in greatly broadening the specificity of SBL is through polarity effects.

Consistent with this observation, S166C-S-j,k were able to catalyze ligations with a similarly broadened specificity. Some small but significant differences in yield were observed for S166C-S(CH2)nNH2+ (j) and S166C-SCH2(C2H4NH3)+3 (k). For example, with S166C-S-j,k as catalysts, hydrophobic P1 substrate 2 gave distinctly lower yields (by 17–18%) of 10.
as compared with SBL-WT. In contrast, catalysis by S166C-S-k, which is modified with a triple positive charge at the base of the S1 pocket, gave the highest yield in the coupling of 4, which bears a negatively charged P1 glutamate side chain, to give 12 in 71% yield. These opposing effects suggest the creation of a more highly charged S1 pocket environment that disfavours the binding of the non-polar benzyl side chain of 2 and favours, perhaps through complementary electrostatic interactions, the binding of 4; an effect that mirrors those found in some kinetic analyses.[5f, i]

Generally, however the increase in charge at position 166 from /C135 to /C1353 had little effect on yield (Table 2) and this discounts a purely electrostatic mechanism.

The need for polar side chains at position 166 to create the most effectively broadened catalysts was further confirmed by the use of S166C-S-l bearing a non-polar yet similarly bulky position 166 side chain. Although able to catalyse the synthesis of 10±13 in yields approaching those of polar CMMs S166C-S-a±k, yields for the more challenging syntheses of 14±18 were either zero or substantially inferior. In displaying this lack of breadth, S166C-S-I with its less polar S1 pocket behaves more like SBL-WT than broad-specificity polar CMMs S166C-S-a±k.

The use of CMMs in trial library syntheses: Valuably, the greatly broadened specificity of these CMMs and therefore their potential as combinatorial biocatalysts[5, 3] was demonstrated by the parallel synthesis of dipeptides of L-Phe, L-Ala, L-Glu, d-Ala and d-Glu to form a small array (Scheme 3). CMMs S166C-g,j,k were selected on the basis of their success in individual ligations. These were compared with SBL-WT as a reference unbroadened catalyst.

Thus, each of the four catalysts, S166C-g,j,k and SBL-WT, under identical conditions (1:1 DMF:water, 48 h) were used to transform the following reaction pairs: 2 + 8, 3 + 8, 4 + 8, 6 + 8, 7 + 8, 4 + 9 to create a small 24-member array of dipeptides 10, 11, 12, 15, 17 and 18.

Examination of Scheme 3 highlights that SBL-WT (Row 4, Catalyst Array) is clearly incapable of creating libraries of dipeptides in this way. We were therefore delighted to find that all three CMMs S166C-g,j,k yielded their dipeptide products in >70% yield (71%–99%), as judged by multowell LC-ESMS analysis. This balanced, high yield of each of 10, 11, 12, 15, 17 and 18 mirrors, and in some cases exceeds, the results obtained in our initial evaluations of synthetic utility.

Conclusion

On the basis of positive indications in screens of enzyme activity and specificity and through evaluation in preparative-scale peptide ligations, we have shown that the introduction of polar substituents (a "polar patch") to the S1 pocket, such as carbohydrates (a–d), oxazolidinones (e–i) and alkylammonium groups bearing up to three positive charges (j, k) creates beneficial electrostatic and/or hydrogen bonding interactions to greatly broaden the synthetic substrate specificity of serine protease SBL.

Dramatically all of the CMMs S166C-S-a–k were able to catalyze the coupling of all of the d-amino acid donors 5–7 with acyl acceptor 8; reactions that SBL-WT cannot perform. Indeed in some cases the natural L-preference is reversed, for example, based on yield, S166C-S-j showed a 1:2:1 stereochemical preference for d-glutamate 7 over L-glutamate 4. In 1997 the bold statement was made that "the efficiency of d-peptide formation cannot exceed that of L-
Remote effects of these position 166 S modifications upon the neighbouring S₁ pocket were also observed. These allowed the syntheses of C-terminal l-alaninyl dipeptides 14 and 15 that again are not synthesized by SBL-WT. Interestingly, however, all CMMs S166C-S-a-I gave lower yields of 13 than SBL-WT. This intriguing single exception to the enhancement of synthetic utility though the “polar patch” CMM approach may be due to an indirect reduction in the volume of the already small S₁ pocket. From inspection of Figure 1 it can be seen that residues 154–156 in the wall of the S₁ pocket encroach on the wall of the S₁ pocket. Any bulging or distortion in this wall might therefore restrict the S₁ pocket. This is consistent with the observed drop in yield of 13 (Table 2) as the size of the position 166 side chain increases:

-CH₂OH (57 %) → CH₂S-j,j,1 (42–50 %) → CH₂S-a–h,k (28–37 %).

We have also established that a CMM bearing a non-polar modification, that is, S166C-S-I, at position 166 is far less successful in these ligations. We therefore speculate that “polar patch” modifications alter the nature of the specificity of the S₁ pocket or even disrupt its role by exposing it to solvent: perhaps effectively turning the 152-167 loop of SBL into a more flexible, scaffold. [43] This is consistent with previous observations by Dordick and co-workers that a polar mutation at position 166 (G166N) in the closely related enzyme subtilisin BPN′ significantly enhanced activity through transition state stabilisation, particularly with polar P₁ substrates.[44] Perhaps as a consequence, this increased binding pocket flexibility broadens the range of side chains that SBL can accommodate in this pocket.

Other factors cannot be excluded. For example, complementary electrostatic interactions between the side chain carboxylate of substrate suc-AAPePNA and a CMM S166C-S-(CH₂)₂NH₃⁺ resulted in an 19-fold increase in kcat/Km over SBL-WT[56] and hydrogen bonding interactions between substrate and the ß-glucoside residue of a CMM L217C-S-Glc(Ac), resulted in an 8.4-fold increase in esterase activity.[51] Perhaps strong polar interactions between, for example, the carboxylate side chain of 4 and the side chain of the S166C site of, for example, CMMs S166C-S-j,k provide a more stable ES-complex, which cannot be easily attacked by water, thereby allowing preparation of dipeptides in good yield. Moreover, these CMMs may even bind ß-amino acids in a different mode from 1-amino acids, that is, the ß-carbonylveroximegroup (NHZ) of may bind in the S₁ pocket in place of the amino acid site of the CMMs could cause the low reactivity that was observed for 5. Any or all of these factors may play a role in the broadening that has been observed. Further studies to explore these interesting possibilities are in progress, however from this work the required effects for successful broadening may be tentatively ranked polarity > size > charge.

It should be noted that although many of these yields are obtained over different reaction times, the use of CMMs in initial trial parallel syntheses under identical conditions proved highly successful. This is a first step towards the combinatorial biocatalysts that we require. It is also clear from our work that although kinetic analysis is a valuable screening tool in the identification of candidate biocatalysts, this should be coupled with the type of broad ranging evaluation of synthetic utility described in this paper.

In conclusion, we have established a significant enhancement of the applicability of SBL in peptide synthesis using “polar patch” CMM technology. We have achieved our goal of creating, SBL-CMMs S166C-S-a–k that accept a wide range of substrates including ß-amino acids as acyl donors and an ß-branched acyl acceptor to give a variety of dipeptides, many in very high yields, that cannot be synthesized with SBL-WT. Furthermore, these dramatic improvements have been achieved without the loss of the natural specificity of SBL. Resulting “polar patch” CMMs are consequently successful in creating balanced dipeptide libraries that SBL-WT cannot. These results therefore represent a true broadening of synthetic utility. This methodology is a powerful tool for enhancing the application of enzymes in organic synthesis. Initial work on broadening the specificity of other synthetically useful enzyme systems has already begun.[45]

**Experimental Section**

General: ¹H and ¹³C NMR spectra were measured on a Varian Unity 400 (400 MHz for ¹H and 100 MHz for ¹³C) spectrometer in [D₆]DMSO with residual solvent as internal standard. High-resolution mass spectra (HRMS) were recorded using a Micromass ZAB-SE spectrometer (FAB+). Protein mass spectra were recorded using a Micromass Platform or LCT spectrometer (ES⁺) and deconvoluted by using MaxEnt. Optical rotations were measured with a Perkin-Elmer 243B polarimeter. ALUGRAM SIL G/UV254 Art.-Nr. 818 133 (Macherey-Nagel) was used for analytical TLC. Preparative TLC was performed on a pre-coated Silica gel plate Art.5744 (Merek). Amino acids and derivatives were purchased from Sigma or Bachem and were used as received. All solvents were reagent grade and distilled prior to use.

**General CMM preparation procedure**: Mutants of Subtilisin Bacillus lentsus (SBL) were generated, and WT and mutant enzymes purified as described previously.[5a,c] A frozen aliquot of the mutant enzyme S166C (containing from 25 mg to 1.1 g of enzyme at a concentration of approx 20 mg mL⁻¹) was thawed and added to an equal volume of Modifying Buffer (for S166C-S-): 140 mM CHES, 2 mM CaCl₂, pH 9.5; for S166C-S-j,k: 140 mM MES, 2 mM CaCl₂, pH 6.5; for S166C-S-h,k: 140 mM MES, 2 mM CaCl₂). To this solution was added 100 µL of a 0.2 mM MTS reagent solution for every 2.5 µL of the resulting enzyme solution (1a–ce–h,1d in water). The mixture was sealed, vortexed and shaken at room temperature. When completion of modification was determined by a specific activity assay using succinyl-Ala-AlaPro-Phe-γ-nitroanilide (kcat = 8800 s⁻¹ cm⁻¹)[50] as substrate in 0.1 mM Tris-HCl buffer containing 0.005% Tween 80, 1% DMSO, pH 8.6 showing constant activity and by titration with Ellman’s reagent[50] (ε₂₈₅ = 13600 M⁻¹ cm⁻¹) showing no free thiol present in solution. A further solution (50 µL for every 2.5 µL of enzyme solution) of the modifying reagent solution was added and the mixture agitated for a further 10 min. The reaction was poured onto a pre-equilibrated G-25
Sephadex column and eluted with Quencher Buffer (5 mM MES 1 mM CaCl₂, pH 5.5). The eluant was dialysed at 4°C against 10 mM MES, 1 mM CaCl₂, pH 5.2 (2 × 45 min). The resulting dialysate was flash frozen in liquid nitrogen and stored at −18°C. The free thiol content of all CMMs, was determined spectrophotometrically by titration with Ellman’s reagent[20] in phosphate buffer 0.25 M, pH 8.0. In all cases no free thiol was detected. Modified enzymes were analysed by nondenaturing gradient (8–25%) gels at pH 4.2, run towards the cathode, on the Pharmacia Phast-system and appeared as single bands. Prior to ES-MS analysis CMMs were purified by FPLC (BioRad, Biologic System) on a Source 15 RPC matrix (17–0727–20 from Pharmacia) with 5 % acetonitrile, 0.11% TFA as the running buffer and eluted with 80% acetonitrile, 0.11% TFA in a one step gradient.

EMSS (MS-MS) m/z: S166C-S= calculated 27036, found 27037; [M + H]+ 379.42; 280.1297, found 280.1303.


Z-S-AlaGlyNH₂ (14): [m/z] = 218.09 (2.0 μM in MeOH) [25]; HRMS (ES): Found 218.0906; calcd for C₁₃H₁₈N₃O₄ [M + H]+: 218.1256.


Z-S-PhGlyNH₂ (16): [m/z] = 3.4 (1.17 μM in MeOH); H¹ and ¹³C NMR spectral data identical to 10. HRMS calculated for C₁₅H₂₀N₃O₆ [M + H]+: 284.1454, found 284.1457.

Z-S-GlyNH₂ (17): [m/z] = 8.5 (0.6 μM in MeOH) [25]; H¹ and ¹³C NMR spectral data identical to 11. HRMS calculated for C₁₅H₂₀N₃O₆ [M + H]+: 280.1297, found 280.1303.

Z-S-GluNH₂ (18): [m/z] = 9.1 (0.6 μM in MeOH); H¹ and ¹³C NMR spectral data identical to 12. HRMS calculated for C₁₅H₂₀N₃O₆ [M + H]+: 338.1352, found 338.1353.

Acknowledgement

We acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC), Genencor International, Inc and the University of Oxford for generous funding. We are also grateful to Genencor International, Inc., for providing WT- and S166C-SBL enzymes, and to Drs Rink Bott and Kanjai Khumvitveeporn for helpful discussions. We also thank the
Engineering and Physical Sciences Research Council of the UK (EPSRC) for access to the Mass Spectrometry Service at Swansea and the Chemical Database Service at Daresbury.


[20] Numbering based on the numbering of subtilisin BPN’.


